LIBRARY
UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

MODERN LANGUAGE

NOTICE: **Return or renew all Library Materials! The *Minimum* Fee for each Lost Book is $50.00.**

The person charging this material is responsible for its return to the library from which it was withdrawn on or before the **Latest Date** stamped below.

Theft, mutilation, and underlining of books are reasons for disciplinary action and may result in dismissal from the University. To renew call Telephone Center, 333-8400

UNIVERSITY OF ILLINOIS LIBRARY AT URBANA-CHAMPAIGN

Modern Lang & Ling Library

UIUC Rm 425 333-0076

JOURNAL

1 WEEK

L161—O-1096

-92
8
v.1

mod. Lang

Studies in The Linguistic Sciences

8, No. 1
ng 1978

PAPERS IN GENERAL LINGUISTICS

ESTHER BENTUR *Orthography and the Formulation of Phonological Rules* 1

PETER COLE, WAYNE HARBERT AND GABRIELLA HERMON *Headless Relative Clauses in Quechua* 26

PETER COLE, WAYNE HARBERT, GABRIELLA HERMON AND S.N. SRIDHAR *On the Acquisition of Subjecthood* 42

PETER COLE AND JANICE JAKE *Accusative Subjects in Imbabura Quechua* 72

JANICE JAKE *Why Dyirbal isn't Ergative at all* 97

YAMUNA KACHRU AND RAJESHWARI PANDHARIPANDE *On Ergativity in Selected South Asian Languages* 111

JAMES LEVINE *The Relative Pronoun and the Long Form Adjective in Russian* 128

DAVID ODDEN *Aspects of Iraqi Arabic Verbal Phonology* 137

RAJESHWARI PANDHARIPANDE *Exceptions and Rule Government: The Case of the Passive Rule in Hindi* 153

IRMENGARD RAUCH *Language-Likeness* 174

PAMELA S. SILVER AND SCOTT R. KRAUSE *A Reanalysis of the Class 5 Prefix in Shona* 181

S. N. SRIDHAR *Linguistic Convergence: Indo-Aryanization of Dravidian Languages* 197

SUSAN U. STUCKY *How a Noun Class System may be Lost: Evidence from Kituba (Lingua Franca Kikongo)* 216

THE LIBRARY OF THE
MAY 09 1979
UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

Department of Linguistics
University of Illinois

STUDIES IN THE LINGUISTIC SCIENCES

PUBLICATION OF THE DEPARTMENT OF LINGUISTICS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

EDITORS: Charles W. Kisseberth, Braj B. Kachru, Jerry L. Morgan

REVIEW EDITORS: Robert N. Kantor and Ladislav Zgusta

EDITORIAL BOARD: Eyamba G. Bokamba, Chin-chuan Cheng, Peter Cole, Georgia M. Green, Hans Henrich Hock, Yamuna Kachru, Henry Kahane, Michael J. Kenstowicz, Chin-W. Kim and Howard Maclay.

AIM: *SLS* is intended as a forum for the presentation of the latest original research by the faculty and especially students of the Department of Linguistics, University of Illinois, Urbana-Champaign. Especially invited papers by scholars not associated with the University of Illinois will also be included.

SPECIAL ISSUES: Since its inception *SLS* has been devoting one issue each year to restricted, specialized topics. A complete list of such special issues is given on the back cover. The following special issues are under preparation: Fall 1978: *Linguistics in the Seventies: Directions and Prospects* (Forum Lectures delivered at the 1978 Linguistic Institute of the Linguistic Society of America), edited by Braj B. Kachru; Fall 1979: *Relational Grammar and Semantics,* edited by Jerry L. Morgan; Fall 1980: *Papers on Diachronic Syntax,* edited by Hans H. Hock; Fall 1981: *Studies in Language Variation: Nonwestern Cast Studies,* edited by Braj B. Kachru; Fall 1982: *South Asian Linguistics: Syntax and Semantics,* edited by Yamuna Kachru.

BOOKS FOR REVIEW: Review copies of books (in duplicate) may be sent to the Review Editors, Studies in the Linguistic Sciences, Department of Linguistics, University of Illinois, Urbana, Illinois, 61801.

SUBSCRIPTION: There will be two issues during the academic year. Requests for subscriptions should be addressed to *SLS* Subscriptions, Department of Linguistics, 4088 Foreign Languages Building, University of Illinois, Urbana, Illinois, 61801.

Price: $5.00 (per issue)

S792
V.8

STUDIES IN THE LINGUISTIC SCIENCES

EDITORS

Charles W. Kisseberth
Braj B. Kachru, Jerry L. Morgan

REVIEW EDITORS

Robert N. Kantor and Ladislav Zgusta

EDITORIAL BOARD

Eyamba G. Bokamba, Chin-chuan Cheng, Peter Cole,
Georgia M. Green, Hans Henrick Hock, Yamuna Kachru, Henry Kahane,
Michael J. Kenstowicz, Chin-W. Kim, and Howard Maclay.

VOLUME 8, NUMBER 1
SPRING, 1978

DEPARTMENT OF LINGUISTICS, UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS 61801

TABLE OF CONTENTS

Esther Bentur: Orthography and the Formulation of Phonological Rules. 1

Peter Cole, Wayne Harbert and Gabriella Hermon: Headless Relative Clauses in Quechua 26

Peter Cole, Wayne Harbert, Gabriella Hermon and Shikaripur Sridhar: On the Acquisition of Subjecthood. 42

Peter Cole and Janice Jake: Accusative Subjects in Imbabura Quechua. 72

Janice Jake: Why Dyirbal isn't Ergative at all 97

Yamuna Kachru and Rajeshwari Pandharipande: On Ergativity in Selected South Asian Languages 111

James Levine: The Relative Pronoun and the Long Form Adjective in Russian 127

David Odden: Aspects of Iraqi Arabic Verbal Phonology. 137

Rajeshwari Pandharipande: Exceptions and Rule Government: The Case of the Passive Rule in Hindi. 153

Irmengard Rauch: Language-Likeness 174

Pamela S. Silver and Scott R. Krause: A Reanalysis of the Class 5 Prefix in Shona. 181

S.N. Sridhar: Linguistic Convergence: Indo-Aryanization of Dravidian Languages. 197

Susan U. Stucky: How a Noun Class System may be Lost: Evidence from Kituba (Lingua Franca Kikongo). 217

Studies in the Linguistic Sciences
Volume 8, Number 1, Spring 1978

ORTHOGRAPHY AND THE FORMULATION OF PHONOLOGICAL RULES

Esther Bentur

Within the framework of generative phonology, a 'rule' stands for an hypothesis about a linguistic generalization perceived by speakers. In a linguistic model which aims at characterizing linguistic competence realistically, only valid generalizations--namely, only those which speakers may abstract given the input data--should be expressed. An adequate analysis should neither overstate nor underestimate the speaker's internalized knowledge. It is usually assumed that only oral data counts as the basis on which hypotheses are constructed by speakers (and consequently by linguists). To demonstrate that this tenet cannot be maintained it will be shown that a generalization, which is recoverable synchronically only from the written data, is nevertheless perceived by speakers and becomes an integral part of their linguistic knowledge.

The example will be constructed as follows:

First, the relevant data and the historical background will be presented. An abstract analysis to account for those facts will be proposed. Several arguments against it will be provided, and consequently, it will be rejected. An alternative proposal will be offered and its validity will be tested in two productivity tests; one involving hypothetical lexical items and the other the elicitation of non-existing extensions of actual lexical items. The results of the second test will point, however, to the inadequacy of the alternative (i.e. non-abstract) proposal. A solution to the apparent paradox will be offered by incorporating an additional factor into the model, namely the influence of orthography. To substantiate the claim, that this is indeed the crucial factor, the results of a third productivity test, this time a written one, will be mentioned.

1. DATA

Consider the patterns of past and present forms of the following verbs:

(1)	kara	'he read'	kore	'he reads'
	kafa	'he froze'	kofe	'he freezes'
	maca	'he found'	moce	'he finds'
	but:			
(2)	kara	'he tore'	korea	'he tears'
	mana	'he prevents'	monea	'he prevents'
	šama	'he heard'	šomea	'he hears'

Whereas the past tense forms of both 'find' and 'hear' have the same configuration CaCa, the corresponding present tense forms exhibit different patterns CoCe (for 'finds', 'freezes' and 'reads') but CoCea (for 'tears',

'prevents' and 'hears'). Comparison with the pattern of regular verbs (i.e. where all the three radicals surface) such as:

(3)	lamad	'he studied'	lomed 'he studies'
	šavar	'he broke'	sover 'he breaks'

suggests, that the present tense singular form in this verb paradigm is $C_1oC_2eC_3$. The CoCea forms, thus, manifest an additional final vowel a (beside the non-occurrence of the third radical).

A similar pattern shows up in the following verbs:

(4)	mašax	'he pulled'	mošex 'he pulls'
	našax	'he bit'	nošex 'he bites'

but:

(5)	barax	'he ran away'	boreax 'he runs away'
	carax	'he screamed'	coreax 'he screams'

Here, the present tense forms are either of the type C_1oC_2ex (as in 'pulls' and 'bites'), or C_1oC_2eax, with an additional [a] preceding the third radical [x] (as in 'runs away' and 'screams').

2. HISTORICAL BACKGROUND

These synchronic patterns of alternations are reflexes of an historical rule which inserted the vowel a when either /h/, /ḥ/ or /ᶜ/ occurred in final position following a non-low vowel (/i/, /e/, /o/ or /u/).

Expressed in distinctive features, the rule can be formulated as:

a-Insertion Rule

(6) $$\emptyset \longrightarrow \begin{bmatrix} +\text{syll} \\ +\text{low} \end{bmatrix} \Big/ \begin{bmatrix} +\text{syll} \\ -\text{low} \end{bmatrix} \_\_\_\_\_ \begin{matrix} C \\ \begin{bmatrix} +\text{low} \\ +\text{continuant} \end{bmatrix} \end{matrix}$$

In MH, however, the following changes have taken place:

(7) /ᶜ/ is always realized as [ʔ] or Ø.

/ḥ/ surfaces always as [x] and thus has merged completely with [x] which is an allophone of /k/.[1]

/h/ is unstable in informal speech and is realized by most speakers as [ʔ] or as Ø.

In spite of those mergers, a still surfaces in MH only when the following consonant was historically /ᶜ/, /ḥ/ and /h/.

3. THE ABSTRACT ANALYSIS

Consider first the analysis which recapitulates the historical events. According to it the a-Insertion Rule would still be a synchronic rule. Its application must precede the merging rules:

(8) $^{c} \longrightarrow$?
$h \longrightarrow x$
$\dot{h} \longrightarrow$?

The ordering is necessary to ensure that [a] will surface only when it is followed by [ʔ] or [x], which are realizations of historical /^{c}/, /$\dot{h}$/ and /h/.

This analysis, which is crucially based on the postulation of underlying /^{c}/, /$\dot{h}$/ and /h/ as phonemes of MH, is maintained by most standard Hebrew grammar books, and was formulated within a generative framework by Ornan (1972). It "explains" the fact that phonetically identical segments behave differently with respect to the application of the same phonological rule. As Ornan mentions, given this analysis, the morphology of Hebrew has become considerably easier.

The price, however, one should note, is a highly abstract phonemic system with underlying segments that never show up on the surface. In what follows, arguments against this abstract analysis will be presented. It will be demonstrated that it cannot possibly be maintained, given the existing internal evidence which clearly points to its inadequacy.

4. ARGUMENTS AGAINST THE ABSTRACT ANALYSIS

The postulation of non-occurring surface segments as underlying elements in the synchronic system has been a controversial issue within generative phonology.[2] Even the proponents of abstract analyses admit that the burden of proof, in such cases, is on the one who advocates abstractness. For many linguists, however, analyses incorporating hypothetical segments are automatically excluded from the range of viable synchronic analyses (see Hooper 1976). No additional arguments are needed to convince these linguists of the inadequacy of the analysis proposed so far, since it involves the positing of the pharyngeals and voiceless glottal fricative as underlying phonemes in MH even though these elements are not manifested on the surface.

Yet, I would like to discuss additional evidence to indicate that an abstract approach in this case of MH cannot be maintained even by those who allow for hypothetical segments under certain conditions (if, for example, several rules crucially refer to them; see Kiparsky 1971).

4.1 The selection of the feature [+low]

It is widely accepted that the advocate of an abstract approach has to demonstrate clearly that the nature of the postulated segment is determined uniquely by the structure of the language. The feature assigned to it must be well-established within the internal system.

In the Yawelmani case (Kisseberth 1969), for example, the abstract segments were assigned the feature [+high]. The selection of this specific property was motivated by the behavior of the corresponding surface segments with respect to several phonological rules. Kisseberth notes that "the choice of u: and i: as underlying source of some surface o: and all surface e: is motivated since no other choice would perform the same function." Arguments along the same lines are provided by Brame (1972) who argues for an underlying voiced pharyngeal fricative in Maltese.

The selection of the feature [+low] for the abstract segments in Hebrew cannot, however, be supported on internal grounds. There are no additional [+low] segments, which trigger the a-Insertion Rule, that could be cited as independent evidence that the presumed underlying feature involved is [+low]; the only (+low) consonants, which historically conditioned the rule, are precisely those whose existence as elements in the synchronic system is being questioned. The assignment of the feature [+low] to those abstract segments is therefore totally arbitrary from a synchronic point of view and cannot be motivated given the internal organization and structure of the language; theoretically other phonemic features could have been posited as well.

4.2 Current changes in MH

The classical argument in favor of the postulation of abstract underlying segment is based on the observation that this specific segment behaves, with respect to several phonological rules, not in accordance with its surface properties, but rather similarly to other surface segments, which share a specific feature. In the Yawelmani case, for example, (Kisseberth 1969), surface long non-high vowels were posited as high ones underlyingly, since they behave like surface high vowels rather than non-high ones. If it can be shown, however, that, with regards to some of the relevant rules, the abstract segment stops functioning as predicted by its presumed underlying features, and starts behaving according to its surface properties, this would cast doubt at the validity of an abstract analysis.

An abstract analysis implies that speakers are able to maintain a phonemic system which includes non-surfacing segments. No subsequent changes in the direction of a more surface-oriented system should be anticipated. Ideally, all rules should continue referring to the underlying (abstract) rather than surface properties. In what follows several developments which have taken place in MH will be discussed; each of them has affected the abstract segments. What is common to all of these changes is the fact that the abstract segments are being treated according to their surface features. A proponent of the abstract analysis would probably be able to provide an explanation for the emergence of each individual development; on the whole, however, he would have to claim that the fact that all of these developments affect specifically the abstract segments is totally accidental.

A non-abstract analysis (i.e. one which maintains that pharyngeals cannot be posited underlyingly in MH) predicts that those developments will inevitably take place. Such changes are not regarded as random individual instances but rather as related developments within a general scheme. Only a non-abstract analysis can thus offer a unitary explanation for the emergence of those changes in MH; the abstract one will have to treat them on a fragmentary accidental basis.

Barkai (1972) mentions several rules which were conditioned historically by /ʔ/, /ʿ/, /ḥ/ and /h/ (henceforth: guttural segments). He argues, that, since historical /ḥ/ still functions synchronically as a guttural with respect to these rules it must be posited underlyingly as such, even though it always shows up on the surface as a voiceless velar fricative [x] (i.e. as a segment which is not [+low]). Otherwise, he states, each of those rules would have to be complicated.

However, for most of those rules mentioned by Barkai, it can be demonstrated that they are no longer obligatory for historical /ḥ/.

4.2.1 The Vowel Lowering Rule

Consider, for example, the synchronic status of the historical Vowel Lowering Rule:

Vowel Lowering

(9) $\begin{bmatrix} +\text{syllabic} \\ \alpha\ \text{high} \end{bmatrix} \quad \begin{bmatrix} -\text{high} \\ -\alpha\ \text{low} \end{bmatrix} \Big/ \underline{\quad} \begin{bmatrix} -\text{syllabic} \\ +\text{low} \end{bmatrix} \begin{bmatrix} -\text{syllabic} \end{bmatrix}$

The reflexes of this rule still show up in MH. Compare verbs in the hiCCiC paradigm which do not have an historical guttural as a first radical,

(10) hixmin 'invite'
hisbir 'explain'

with those having historical /ʔ/ or /ʿ/ in the same position.

(11) heʔezin 'listen'
heʔenik 'grant' (historically: heʿenik)

(The vowel following the guttural is an inserted echo vowel.)

The rule no longer applies obligatorily when the first radical was an historical /ḥ/; variants with a non-lowered prefix vowel have emerged, i.e.

(12) hexvir ∿ hixvir 'become pale' (historically: heḥvir)
hexzir ∿ hixzir 'return s.th. (historically: heḥzir)

Both alternants are found in free variation in the speech of the same individual.

Similarly one finds:

(13) texanek ∿ tixanek 'she / you will suffocate'

texatex ∿ tixatex 'she / you will be cut'

Given an abstract analysis, these data can be accounted for by assuming that the Vowel Lowering Rule has become optional for some mysterious reason.

Thus, if it does apply to underlying /ḥ/ the output, following the merger ḥ x will be [hexvir]. If, on the other hand, Vowel Lowering was skipped in the course of the derivation the output would be [hixvir]. The fact that the rule has become optional in MH is viewed, within this framework, as an accident.

A non-abstract analysis predicts that a variant such as [hixvir] will indeed emerge, since the element which triggered the Lowering rule historically (i.e., /ḥ/ is no longer a segment in the synchronic system. [hexvir], on the other hand, would be treated as an historical residue.

Another related development, which has taken place in MH, is the emergence of the following variation:

(14)	hixtiv ∿ hextiv	'dictate'	(cf. katav 'write')
	hixšil ∿ hexšil	'fail s.o'	(cf. kišalon 'failure')
	hixbid ∿ hexbid	'make heavier'	(cf. kaved 'heavy')

Whereas in the case of hixzir and hixvir the velar fricative surfaces in all related forms (cf. xazar 'return', xazara 'rehearsal', xiver 'pale') there is a k ∿ x alternation in the pairs of the type hixtiv-katav, hixšil-kišalon. To capture the relatedness between the latter pairs, an underlying /k/ must be postulated as the first radical; it surfaces as [x] due to the operation of spirantization. Since the segment which underlies the surface [x] in those verbs cannot possibly be /ḥ/, the occurrence of the alternant with the lowered prefix vowel (i.e. hextiv) cannot be attributed to the effect of a following [+low] consonant. The emergence of these variants may be viewed as a case of analogy.

(15) hixvir - hexvir
hixtiv - X = hextiv

The Lowering Rule has thus been extended to all surface [x]'s regardless of their historical origin. Notice that as a consequence of the merger ḥ x this rule has become a minor one; it applies only to some [x]'s (i.e. the reflexes of historical /ḥ/'s) which had to be lexically marked. The overgeneralization of the rule implies the elimination of the arbitrary marking, a development which results, of course, in grammar simplification. Given the emergence of the variations in (14) it becomes evident why the existence of alternants with the lowered vowel cannot be cited as an argument to support the selection of the feature [+low] to the abstract segment.

4.2.2 Historical constraints on surface strings $\#C\ \overset{C}{[+low]}$, $\overset{C}{[+low]}\ C$

Historically, Hebrew had restrictions on surface strings of the type * $\#C\ \overset{C}{[+low]}$ and * $\overset{C}{[+low]}\ C$; to break such impermissable sequences an epenthetic vowel was inserted (for details see Barkai 1972, Bolozky 1972). The reflexes of these constraints are still manifested in MH:

Compare:

(16) tmixa 'support'
švita 'strike'

with

(17) te?ima 'tasty' (f. sing.)
se?iva 'pumping'

similarly:

(18) mamšix 'continue' (m. sing. pres.)
matxil 'start' (m. sing. pres.)

versus

(19) ma?amin (∿mamin) 'believe' (m. sing. pres.)
ma?atik (∿matik) 'copy' (m. sing. pres.)

These constraints no longer hold for surface [x]'s which are reflexes of historical /ḥ/'s.

Thus we find in MH:

(20) txina 'grinding'
šxita 'slaughter' (compare (17))
taxšov 'you will think'
maxzik 'hold' (compare (19))

These are clear cases where the presumed abstract segment no longer behaves according to its abstract underlying properties, but rather according to its surface features.

Within an abstract analysis this development can be handled by ordering the relevant rules as:

1) ḥ ⟶ x
2) Vowel Epenethsis

A proponent of this approach may even claim, perhaps, that the mere fact that an abstract segment stops functioning according to its underlying properties with respect to one (or even several rules) is an immaterial accident and cannot be regarded as a strong argument against abstractness. He might go on to suggest that as long as for some other rules (or even one??) the segment in question continues to behave in accordance with its abstract properties this would be a sufficient indication for its existence.

A non-abstract analysis, on the other hand, predicts the emergence of such changes, i.e. that rules and constraints will become sensitive to surface features rather than to abstract underlying ones.

4.2.3 The convergence of the le?eCoC and la?aCoC patterns

Another process which is presently taking place in MH supports the claim that the historical phonemic contrast /ʔ/ : /ᶜ/ is being abolished.

Consider the following data:

The infinitive of the Hebrew verb paradigm $C_1aC_2aC_3$ has the configuration $liC_1C_2oC_3$ as in:

(21) lilmod 'to study' (root L-M-D)
lišbor 'to break' (root Š-B-R)

For verbs which had as a first radical /ʔ/, the infinitive surfaced as $leʔeC_2oC_3$; if, however, the first radical was /ᶜ/ the infinitive was of the pattern $laᶜaC_2oC_3$. Compare:

(22) leʔexol 'to eat' (ʔ-K-L)
leʔesof 'to collect', 'to gather' (ʔ-S-P)

but:

(23) laᶜavod 'to work' (ᶜ-B-D)
laᶜamod 'to stand' (ᶜ-M-D)

If the phonemic distinction /ᶜ/ : /ʔ/ is indeed kept in MH one would have expected those patterns to be maintained; this however is not the case. We find the following forms in free variation in standard MH even in the speech of educated adults:

(24) leʔesof ∿ laʔasof 'to collect'
leʔenox ∿ laʔanos 'to rape'
leʔefot ∿ laʔafot 'to bake'
leʔeroz ∿ laʔaroz 'to pack'

In all of the forms in (24) the first radical is historical /ʔ/. Yet it seems that the pattern $laʔaC_2oC_3$ is more common in these cases (only two verbs: leʔexol 'to eat' and leʔehov 'to love' do not show these variations). Within the framework of an abstract analysis which maintains the /ʔ/ : /ᶜ/ contrast, this development must be treated as accidental.

Viewed, however, in combination with other cases which have been cited above, this process is readily explained; the speakers' inability to keep the two patterns apart is another indication that the relevant phonemic contrast (in this case /ʔ/ : /ᶜ/) is no longer maintained in their phonemic system.

4.2.4 The behavior of morphemes containing several abstract segments

Consider the case of lexical items which contain several occurrences of the same presumed underlying segment.

The adjective taxuax 'crumbled, crushed (soil)' is derived from the historical root T-Ḥ-Ḥ. Within an abstract analysis this morpheme would be represented underlyingly as /taḥuḥ/. The third radical behaves as predicted by this analysis; it triggers the a-Insertion Rule. The second radical, however, does not behave as a [+low] segment; the corresponding fem. sing. form in [txuxa] namely a vowel is not inserted to break the impermissible sequence #C $\overset{C}{[+low]}$.

An abstract analysis cannot provide a plausible explanation to the question: Why did this root receive such a mixed representation? Why was the same segment analyzed as [+low] only in one of its occurrences in the same lexical item?

A non-abstract approach, on the other hand, would assume that both the second and the third radical are underlyingly /x/, but one of them (i.e. the third) is arbitrarily marked for the a-Insertion Rule.

4.3 Children's behavior

The linguistic behavior of children speaking MH can be cited as another piece of evidence for the claim that the abstract analysis is inadequate for this dialect of Hebrew.

Verbs whose third radical is historical /ḥ/ or /^{c}/ may surface in children's speech with or without the preceding [a]. On the other hand, there are many cases where overgeneralization of the rule takes place and it applies in verbs whose third radical is historical /ʔ/ or /k/ (remember that historical /k/ as a third radical always surfaces as [x] due to spirantization). There does not seem to be a general principle governing the direction of the process. Bar Adon (1959) concludes that "roots having as their third radical historical /k/ and /ḥ/ got so mixed up and it is rather difficult to distinguish between them. Analogy operates in both directions." He cites the following variations (radicals indicated in accordance with the historical pronunciation).

(25) maxbi ∿ maxbia 'he hides' (H-B-ʔ)
ikpi ∿ ikpia 'he froze' (K-P-ʔ)
imci ∿ imcia 'he invented' (M-C-ʔ)

as well as

cove ∿ covea 'he paints' (C-B-^{C})
bole ∿ bolea 'he swallows' (B-L-^{C})

and similarly

(26) linšox ∿ linšoax 'to bite' (N-Š-K)
lišpoax ∿ lišpox 'to spill' (Š-P-K)

Why is it that the variability occurs precisely with the presumed abstract segments?

An abstract analysis can "explain" this by invoking a principle which states that the acquisition of abstract segments is inherently more difficult and thus will be accomplished only at a later stage. Given this approach, the specific developments which have taken place in MH are to be attributed to children's unfamiliarity with the abstract segments. Once those segments are acquired, however, one would have probably anticipated the elimination of variability and confusion. If this were indeed the case, major revision in adults' grammar should have been expected. Namely, the distinction between infinitives for verbs with first radical /ʔ/ and /^{c}/

would reemerge; Vowel Lowering would become again an obligatory rule. However, we do not find such revisions taking place.

4.4 Conclusions

Each of the separate developments cited above can be accounted for in some way within an abstract approach, and thus when treated individually they cannot be accepted as conclusive arguments against abstractness. Yet, when viewed in combination, these developments constitute strong circumstantial evidence against the abstract analysis.

Notice that the abstract analysis has been rejected not only on the basis of axiomatic beliefs such as "all phonological abstractness in the form of imaginary or abstract segments must be ruled out" (as a consequence of the True Generalization Condition, see Hooper 1976), or because we "take it as self-evident that an acquisition theory faces an extraordinary burden if it is required to account for the learning of the existence and locus of occurrences of phonomes that do not exist as perceptual realities anywhere in the language" (Brian 1974). Rather, it has been rejected since it was proven to be inadequate on internal grounds.[3,4]

The issue of the behavior of historical gutturals in MH has to be viewed given the background of the revival of the Hebrew language. MH has been revived on the basis of Biblical and post-Biblical texts. The Biblical corpus reflects a certain phonemic system in which gutturals functioned as a distinct group and conditioned phonological rules which were phonetically perfectly motivated (Vowel Lowering, for example). The phonemic system of MH is not identical to that of Biblical Hebrew; mergers have taken place and, as a result, phonemic contrasts, on which the application of those phonological rules depended crucially, have been neutralized. Yet due to the relatively short period of time (MH was revived at the beginning of this century) as well as the normativistic pressure exerted by the school system, many surface systematicies, which are rooted in Biblical Hebrew, are still maintained.

Some of the surface systematicies can no longer be treated as the outcome of phonological rules, which recapitulate the historical ones. Rather, some instances have been reanalyzed as morphologically conditioned, while others must be treated as involving 'exception features'.

Given the above mentioned arguments one must conclude that the fact that adult speakers of MH still maintain the a only when the final radical was a historical /ᶜ/, /ḥ/ or /h/ cannot possibly be accounted for by a synchronic rule, which reflects exactly the historical one, since such a rule must necessarily be sensitive to phonemic contrasts which cannot be posited underlyingly in the synchronic system.

5. THE AD HOC SUBCATEGORIZATION ANALYSIS[5]

As an alternative one might consider an analysis according to which morphemes are marked in an arbitrary fashion by a diacritic which will trigger the a-Insertion Rule. The rule itself does not recapitulate the historical rule; it no longer makes reference to [+low] as the conditioning

environment. The question whether a certain morpheme would have an allomorph in which the a surfaces is treated under this analysis as an idiosyncratic property of this morpheme, and is no longer entirely dependent on its phonological make-up. In other words: this fact has to be memorized by the speaker and is not totally predictable on phonological grounds.[6]

To test the adequacy of this analysis two productivity tests were conducted. Before presenting the results, the relevance of such tests for the issue involved will be discussed.

6. THE RELEVANCE OF PRODUCTIVITY TESTS

Kiparsky (1974) suggests that productivity should be viewed as a gradient phenomenon rather than a clear-cut dichotomy between productive and non-productive. On one side of the scale there are alternations which are fully productive in the sense that they extend to every new formation that meets their structural description. Such a productivity pattern would be taken as an indication for the extence of a major rule. On the other end one finds the traditional suppletive patterns of the type go-went, is-are: these would never be productive.

In between are alternations which are only partially productive: new words may but do not necessarily have to become subject to them. Berko (1958) found for example that 40% of the adults in her experiment when presented with the word *heaf gave *heaves as the plural rather than *heafs. She remarks:

"When a small group of common words exist as a category by virtue of their great phonetic similarity and their morphological consistency, a new word having the same degree of phonetic similarity may be treated according to this special rule."

These would be cases characterized in standard generative phonology by minor rules (or ad hoc subcategorization). In the productivity pattern of those cases one does find variability among speakers, as well as inconsistent behavior within one speaker, who might apply the rule only in some cases but fail to do it in others, even though they meet the same phonological conditions.

This behavior follows from the fact that the application of those rules depends not only on certain phonological configurations but is determined also by an arbitrary marking, whose existence is unpredictable. In cases involving minor rules and ad hoc subcategorization there will always be two possible outputs, each exhibited by some actual morphemes in the lexicon. When confronted with a new formation, which meets the structural description of the rule, the speaker has to decide whether to mark it as [+] or [-] the rule. This is an arbitrary decision, since nothing in the phonological shape of the new morpheme will dictate the 'correct' output. All the speaker can rely on is his previous acquaintance with lexical items which seem to structurally match the new one. Whether or not the rule will apply to the new formation depends on the pattern of alternation manifested by the actual lexical item selected as a model of reference. The factors determining the association of a new item with a specific model are intricate and may vary

from speaker to speaker. All this predicts that the productivity of such pattern may be random rather than systematic (see Ohala 1973).

If by psychological validity of a rule we mean 'perception of a certain recurring pattern', then rules can be psychologically real without necessarily being extended to all new formations which meet their structural description. The variability factor, mentioned earlier with respect to rules involving ad hoc marking refers only to the question of applicability of a certain rule; once it applies the output is totally predictable. This pattern of productivity clearly differs from that of the traditional suppletive cases: those involve a pattern of alternation which is manifested in one lexical item (is-are, go-went) or by a small class of morphemes which lack a common phonetic denominator (ox-oxen, child-children), and would hardly ever be productive.

It will be assumed therefore that the degree of productivity of phonological rules should be taken not only as a verification of its psychological validity but also as an external manifestation of its most adequate characterization in the speaker's grammar.

The subcategorization analysis suggested for the Hebrew data makes the claim that the application of the rule involves arbitrariness. This approach predicts that in productivity tests variability would be expected. If on the other hand, a systematic behavior will be found, this would be unexplainable by the proposed analysis and, consequently, would point to its inadequacy.

7. EXPERIMENT I - HYPOTHETICAL VERBS

The subjects were 20 native speakers of Hebrew (10 male and 10 female). They were selected from among the Israeli students in Urbana-Champaign who agreed to participate in the experiment. All of them had high school education at least; 12 had academic education as well in either science or engineering, i.e. in areas that do not relate directly to linguistics in general or to Hebrew language in particular.

The experiment involved the formation of present tense forms from the 7 hypothetical past tense verbs. The three root radicals of each of those verbs do not occur in this combination in any actual lexical item.

Each subject was tested individually. The items were presented to the subjects wholly auditorily, one at a time. The two sample items (šamarti - 'I guarded' and katavti - 'I wrote') as well as the two practice items (gamarti - 'I finished' and zaxarti 'I remembered') were 'neutral' in the sense that they did not involve the application of the a-Insertion Rule. Thus the subjects were not deliberately biased in favor of any of the two possible outputs. The subjects were told that the purpose of the experiment was to gather information about the ability of Hebrew speakers to use new verbs. Two lexical items which do not meet the conditions for the a-Insertion Rule, were included to divert the subjects' attention from the real purpose of the experiment.

Table 1 represents the pattern of application of the a-Insertion Rule for each of the relevant items.

Table 1

Responses in the Hypothetical Verbs Test

	+ a Insertion CoCea(x)	- a Insertion CoCe(x)	Discarded
a) ganati	9	12	1
d) lamati	8	10	4
e) pamaxti	18	4	
f) namati	10	12	
g) nalaxti	14	8	

Notice that even though there were 20 subjects, the total responses for each stimulus was 22; this fact will be explained momentarily.

Table 2 represents the pattern of the rule application for each subject (the letters refer back to the stimulus-verbs in Table 1).

Table 2

Pattern of Rule Application for Each Subject

No. of Subjects	+ a Insertion	- a Insertion
5	a, d, e, f, g	_____
2	_____	a, d, e, f, g
4	e, g	a, d, f
3	a, f	d, e, g
4	a, e, f	d, g
2	a, d, e, f, g	a, d, e, f, g

A high degree of variability was manifested in the responses of the group as a whole (Table 1). Furthermore, for half of the subjects there was what seems like a totally inconsistent behavior with regards to the application of the rule; they applied it only for a few items even though all the five hypothetical verbs met the structural description of the rule (Table 2).

A possible explanation for the variability is that the subjects were analogizing to lexical items they already had in their vocabulary. They

were aware of the possible patterns but their choice of each was arbitrary. They might have referred to the pattern of alternation which is manifested by the actual item which was picked as a model. Reference to macati 'I found' or mašaxti 'I pulled' which have present tense forms moce and mošex respectively, would lead to an output such as gone and nolex (i.e. represented as [-'a' insertion]). If on the other hand šamati 'I heard' or baraxti 'I ran away' serve as a model, the outputs would be gonea and noleax respectively by analogy to the present tense forms šomea 'he hears' and boreax 'he runs away'. Since any of these actual lexical items could serve as an adequate model, no wonder that two of the subjects have come up with two possible outputs in each case (that is how we got a total of 22 responses for each stimulus rather than 20).

The results so far seem to support the ad hoc subcategorization analysis.

8. EXPERIMENT II - LEXICAL GAPS

The subjects were the same as in the first experiment.

The second experiment involved the creations of non-existing extensions to actual lexical item (such forms are referred to as 'lexical gaps'). One of the characteristics of Hebrew (and other semitic languages) is the fact that nouns and verbs which are semantically related are usually derived from the same root, i.e. share the consonantal element. Whereas in English one 'draws a picture' and 'wears a sock' in Hebrew he mCaYeR CiYuR and GoReV GeReV respectively (a similar case in English is that of cognate objects, i.e. dream a dream).

A list of 7 Hebrew nouns which do not have verbs derived from the same root as actual lexical items, was compiled and the subjects were asked to come up with related verbs (see Appendix A).

The stimulus-nouns were presented orally one at a time. To facilitate the subjects' task, examples of actual nouns and verbs which manifested the derivational relationships and were semantically close to the stimulus nouns were provided. For example:

'To wear a sock' is ligrov gerev.
'To wear a belt' is laxgor xagora.

How would you say:

He wears a hat? 'a hat' is kova
He puts on a blanket? 'a blanket' is smixa

On the basis of the actual lexical items, which manifest semantic and morphological relationships, and were provided as a model, the subjects were expected to come up with hypothetical verbs which would be related to the stimulus noun in the same fashion (i.e. be derived from the same root).

In the case of masexa and zxuxit, for example, the final radical is realized as [x] ([a] and [it] are suffixes). Since in the infinitive, or

present tense singular masculine, this radical occurs in final position, the structural conditions for the a to be inserted would be met.

The results are indicated in Table 3.

Table 3

Responses in the Lexical Gaps Test

Stimulus		+ a Insertion		- a Insertion		Discarded Responses
kova	'hat'	20	kovea		-	-
smixa	'blanket'		-	20	somex	-
masexa	'mask'		-	20	moxes-8 memasex-12	-
calaxat	'plate'	20	coleax-7 mecaleax-13		-	-
garin	'seed'	11	gorea-4 megarea-7		-	9
emca	'middle'	20	mamcia		-	-
zxuxit	'glass'		-	20	zoxex-7 mezaxex-9 mezaxex-4	-

All of the subjects behave identically with respect to the application of the rule (even though the actual outputs differed depending on the choice of the verb paradigm).[7] At first, it might seem that the results actually support the abstract analysis, which has been earlier rejected (cf. 4). Thus, one might argue that the subjects' systematic behavior was due to the fact that the identical surface segments differed underlyingly, some of them being pharyngeals and that this is why they conditioned a-Insertion. Since the stimulus-nouns were everyday words, speakers have probably had the opportunity to establish these crucial underlying contrasts.

Beside the theoretical arguments and empirical evidence against the validity of an abstract analysis mentioned above, there is an additional reason why such an approach cannot possibly provide a satisfactory explanation in the case of the specific lexical items included in the second experiment. The recoverability of abstract segments crucially depends on the availability of surface clues, namely the presumed abstract segment must behave in a unique way which is not predicted by its surface properties. Its behavior must provide clear indications that its underlying features are indeed distinct from its phonetic makeup.

The stimulus-nouns in the second experiment, however, do not have derivatives in which the third radical of the root occurs in an appropriate

environment to either undergo or condition any rule which could identify it as being underlyingly something different from its surface appearance. Given the non-existence of any clues which could have enabled speakers to differentiate between identical surface segments, it must be concluded, that an abstract approach, even if sometimes it can be motivated for handling data involving alternations, cannot possibly offer a plausible explanation for the subjects' systematic behavior in the second experiment.

The other alternative, namely the arbitrary marking analysis, cannot account for the results either. It is evident that speakers do 'know' more about the pattern involved than what is predicted by such an analysis. If we hypothesize that the subjects were merely analogizing to existing lexical items, we would have expected in such a case to find a great degree of variability, as was found in the first test. A stimulus-noun such as smixa 'blanket', for example, could be analogized on the basis of phonetic similarity, either to cmixa 'growing', slixa 'pardon', 'forgiveness' (which have corresponding verbs comeax and soleax) or to tmixa 'support' nšixa 'a bite' (whose related verbs are tomex and nošex). If the hypothetical verb 'to put on a blanket' was indeed derived by analogy, one would have expected both somex and someax as possible outputs; only the first was, however, suggested by all the subjects.

For other stimuli it was rather difficult to find an actual lexical item, which could serve as a model. For zxuxit a completely matching morpheme does not exist in the whole lexicon of MH. Under such conditions one might have expected the subjects simply to guess the 'correct' output, and as a result to exhibit variability in their responses. Instead, a totally systematic pattern of productivity was found.

A possible explanation might be that speakers were governed in their responses by the tendency to avoid homophony. Thus, since lehamci 'invent' is an already existing item, lehamcia was proposed for 'put in the middle'. This hypothesis, however, can be easily refuted: whereas kovea and somex are actual lexical items (meaning 'decide' and 'rely on' respectively) the other theoretically possible outputs, namely, kove and someax do not occur in the lexicon of Hebrew, yet all of the subjects suggested the former rather than latter.

The only plausible way to account for this systematic linguistic behavior is to conclude, that the subjects were not operating by analogy to subjective models, but rather all of them were applying the same rule: what conditions the application of this rule is, however, a mystery since it cannot be attributed to any surface or underlying difference, as has been demonstrated earlier.

We seem to have arrived at a paradox; earlier it has been argued, that a speaker of MH could not have formulated a major rule to account for the forms where the a surfaces. On the basis of the second test we have come, however, to an opposite conclusion, namely, that speakers do operate with a very general rule.

9. THE RELEVANCE OF ORTHOGRAPHY

Both tests were conducted orally. The second one, however, involved stimuli which were every-day words whose spelling was obviously known to all the subjects.

The outputs suggested in the second test reveal the following pattern:

The a surfaces only when the third radical is spelled with the Hebrew letters 'xet' ח and 'ayin' ע but never where it is spelled with 'kaf' כ and 'alef' א . In other words, it was inserted only in words that were spelled with letters corresponding to the historical /ᶜ/ and /ḥ/. Thus, the phonemic contrasts between /ʔ/ : /ᶜ/, /ḥ/ : /x/, which have been absolutely neutralized in MH, but are still kept in the modern orthography, serve as clues for the modern speaker.

The first test involved non-existing lexical items. Given that their orthographic representations were unknown and could not have been utilized for the purposes of the application of the rule, all the subjects could do was operate by analogy: that is what caused the variability.

To test the validity of this hypothesis an additional test was conducted.

10. EXPERIMENT III - WRITTEN VERSION OF I

The experiment was conducted with the same subjects as in the first and second tests.

The task and material were identical to those of Experiment I, i.e. to form present tense forms of hypothetical past tense verbs.

Each of the hypothetical verbs (cf. 7) was presented orthographically and each subject was asked to pronounce its present tense form. The third radical was indicated half of the time by כ or ח and half of the time as א or ע. The experiment took place a month after the first one.

The results clearly confirm the hypothesis about the use of orthographic clues for the purposes of application of the rule; the [a] surfaced only when the third radical was spelled with ח or ע but never when it was represented as כ or א .

The results of the third test can be cited as additional argument against the validity of an abstract analysis. Such an analysis clearly fails to account for the differences in speakers' behavior in these three tests.

A proponent of the abstract approach might argue that the randomness in the first test has to be attributed to the difficulties in the recoverability of the abstract segments. He would probably try to account for the systematic behavior in the second test by claiming that the lexical representations of these familiar items contain a pharyngeal. Why is it, one must ask, that the responses in the third test differ significantly from those of the first one, although both included the same hypothetical lexical

items? If the systematic behavior in the third test was due to the recoverability of the abstract segments what made this process suddenly possible? Obviously it was not any new phonetic information, since the test was a written one. The only difference between the first and third test was the availability of orthographic representations in the latter. Given this fact, even an advocate of an abstract analysis would have to admit that the orthographic element must be treated as a relevant factor affecting speakers' linguistic behavior.

11. CHILDREN'S RESPONSES

The orthography-based analysis makes the prediction that in the case of preliterate speakers, variability would be found in the second test (the one involving lexical gaps) as well. No significant differences should be expected in their performance in the first and second tests, since the only factor distinguishing between the stimuli in these two tests, namely the orthographic one, would be irrelevant in their case.

To verify this hypothesis the second test was conducted with 15 first graders in an Israeli elementary school. The results are represented in Table 4 (for comparison, the adults' responses to the same stimuli are indicated as well).

Table 4

Responses of Adults and First Graders

in the Lexical Gaps Test

	First Graders		Adults	
	+ a Insertion	- a Insertion	+ a Insertion	- a Insertion
kova	7 (46.7%)	8 (53.3%)	20 (100%)	--
smixa	6 (40%)	9 (60%)	--	20 (100%)
masex	8 (53.3%)	7 (46.7%)	--	20 (100%)
zxuxit	9 (60%)	6 (40%)	--	20 (100%)
calaxat	10 (66%)	5 (33.3%)	20 (100%)	--
emca	9 (60%)	6 (40%)	20 (100%)	--

Nine of the first graders manifested a 'systematic behavior' in the sense that they treated all of the items alike: 4 always inserted the vowel a, the other 5 never did. There were only 2 out of the 15 whose responses were identical to those of the adults.

The uniformity in the adults' performance versus the variability in

the first graders' responses are rather striking. They clearly confirm the hypothesis that the orthographic pattern serves as the crucial clue for the application of the rule.

12. IMPLICATIONS

It seems that as consequence of familiarity with the spelling of lexical items, speakers might realize the correlation between certain orthographic representations (i.e. letters) and a pattern of morphophonemic alternation. Based on the contrast which is still reflected in the orthography the modern speaker is able to formulate a generalization which is rather close to the historical a-Insertion Rule; as has been argued earlier, this rule could not have been recovered independently from the available surface data.

The exposure to orthography might thus lead to modification of the speaker's grammar; patterns of alternations which synchronically involve arbitrariness (i.e. marking with a diacritic) may be reinterpreted as totally predictable when orthographic clues are available. For an educated speaker of Hebrew the contrast between kore 'he reads' and korea 'he tears' is no longer arbitrary; it is tied up with the orthographic distinction (א versus ע) which reflects the historic phonemic contrast.

One might suggest that the spelling pattern serves in this case as a diacritic, i.e. as an arbitrary symbol which triggers the application of the rule.

It seems desirable to avoid using the term 'diacritic' in this context because of its connotations; when used within the current linguistic theory it implies an element which is external and totally ad hoc. Even though orthographic representations are arbitrary to a certain extent, they clearly have psychological reality for educated adults and cannot possibly be regarded as of the same status as a mere diacritic of the type [+rule X] (+rule 235]. One might propose that since orthography is external to the phonological system it should still be referred to as a diacritic. Such a definition is valid only within the basic theoretical assumptions of modern linguistics. The orthographic factor can be treated as external only because the domain of linguistics has been pre-defined as the spoken variety of language alone. Such a division seems to be unmotivated. Rather, the orthographic system should be incorporated and treated as relevant linguistic data, since it clearly has its effect on the formulation of the speaker's grammar.

Within a model which incorporates orthography as legitimate linguistic data available to speakers and utilized by them, the a-Insertion Rule can no longer be viewed as triggered by a totally ad hoc marking. It can be claimed that for literate adults the rule has actually become a major one: it applies whenever its structural description (i.e. a specific grapheme) is met.

The version of the synchronic a-Insertion Rule suggested here, however, is not a regular phonological rule. There is no empirical evidence to support the hypothesis that spelling has indeed affected the speaker's grammar

to such an extent as to introduce new phonemic segments (i.e. [+low] ones). Rather the application of the rule crucially depends on orthographic clues.

The best that can be done is to formulate the synchronic rule by incorporating non-phonological information into its structural description, i.e.:

$$(27)\quad \emptyset \longrightarrow \begin{bmatrix} +\,\text{syll} \\ +\,\text{low} \end{bmatrix} \Big/ \begin{bmatrix} +\,\text{syll} \\ -\,\text{low} \end{bmatrix} \_\_ \begin{Bmatrix} \aleph \\ \cap \end{Bmatrix} \#\#$$

Or, if one accepts the position that historical /^{c}/ are postulated synchronically as /ʔ/ whenever they serve as one of the root radicals (see Bolozky 1972), the a-Insertion Rule can be formulated as follows:

$$(28)\quad \emptyset \longrightarrow \begin{bmatrix} +\,\text{syll} \\ +\,\text{low} \end{bmatrix} \Big/ \begin{bmatrix} +\,\text{syll} \\ -\,\text{low} \end{bmatrix} \_\_ \begin{Bmatrix} ʔ \\ x \end{Bmatrix} \#\#$$

condition: ʔ = ℵ

x = ∩

Such a formulation implies that for the purposes of the rule application, literate speakers of MH rely on orthographic distinctions rather than on phonemic contrasts.

In spite of all these unconventional properties of the rule, it has been clearly demonstrated above that the new formulation reflects rather accurately the speaker's linguistic knowledge. The results of the second test indicate that the orthographically based major rule, rather than the version involving arbitrary marking, is the psychologically valid analysis. Ignoring the relevance of orthography might lead to a distorted picture of what the literate speaker's competence actually is.

The influence of the orthographic system must be viewed as potential rather than automatic. Not every historical rule which is potentially recoverable from the modern orthography will necessarily become part of speaker's grammar: whereas speakers of Hebrew do perceive the historical correlation between the spelling pattern and the application of the a-Insertion Rule, other rules which depended originally on the same phonemic contrasts, and are reflected in the modern orthography as well, have not been maintained with their historical environments but are memorized as arbitrary alternations.[8] This fact does not, however, invalidate the claim made here, namely that historical information may sometimes reach synchrony due to its recoverability from the orthographic system.

To sum up; a common criticism directed towards phonological analyses involving 'abstractness' is that they have lost psychological validity by overstating the speakers' ability to abstract generalizations from given surface data. This view is expressed, for example, by Braine (1974), who notes that the issue is not that such analyses "fail to capture generalizations; rather, it is whether the native speakers make generalizations as the theory captures them."

What I have suggested is that disregarding the relevance of orthographic data in phonological analyses, results again, in a misrepresentation of the speaker's knowledge; this time, however, by understating it; i.e. by failing to express certain valid generalizations. Analyses containing rules referring to contrasts which have been absolutely neutralized on the surface are generally considered as expensive if not automatically ruled out. The results mentioned above suggest, however, that historical rules whose synchronic formulation involves reference to segments which never surface may nevertheless be recapitulated and become an integral part of a speaker's system as long as the relevant historical distinctions are reflected in the orthographic system (assuming of course that we are dealing with a literate society).

NOTES

*The paper is a chapter from a dissertation entitled, "Some Effects of Orthography on the Linguistic Knowledge of Modern Hebrew Speakers." I would like to thank Charles Kisseberth, Hans Hock, and Michael Kenstowicz for the valuable comments and suggestions they provided.

[1]Historically, [x] was an allophone of /k/ derived via spirantization. Historical /k/'s which occur as the third radical, would always surface as [x]'s, because spirantization in Hebrew applied postvocalically, and there is never in the verb paradigm a configuration where a third radical occurs postconsonantally (a position where spirantization would not apply). Thus these [x]'s never alternate with [k]'s. Whether this surface [x] has to be represented synchronically as a phoneme or still as an allophone derived from /k/ is irrelevant for the present purposes. All that matters is that there is no way to tell [x]'s which were historically /ḫ/ from those which were allophones of /k/ in the third position: both are always realized on the surface as [x].

[2]For the controversy about the abstractness of phonological analyses see Kiparsky (1968), Kisseberth (1969), and Crothers (1971).

[3]Bolozky (1972) suggests that gutturals should be still postulated underlyingly primarily because of pattern correspondences. He argues "If one wants to account for awareness of the connection between root, pattern and semantic reading, and for semantic correspondences between realizations of different verbs in the same pattern - then one is logically led to assume gutturals in the underlying representations. If we want yaavod 'he will work', for instance, to be related to avad 'he worked' in the same way that yikšor 'he will tie' is related to kašar 'he tied' the most logical way of doing it is to represent yaavod as /yi+ᶜvod/ and let 'a-insertion' for gutturals and assimilation across a prefix map it into ya?avod (ultimately yaavod)."

Given, however, that those gutturals are hardly ever realized as such on the surface he suggests (p. 81) that "it is quite possible that instead of having /ʔ/, /ᶜ/ and /h/ it would be sufficient to fill the slot they used to occupy by /ʔ/'s only, since it is almost only as [ʔ] that they are

optionally realized. ...Preserving /ʔ/ in the ex-gutturals positions will not cause the loss of significant generalization, and would still satisfy the need for some additional consonants for pattern correspondences and for a plausible explanation of some phonological phenomena attributed to the gutturals. And since it is (almost) only /ʔ/ that is optionally found in the phonetic output our underlying representation would not be too abstract."

The arguments in favor of the postulation of underlying /ʔ/'s do not hold for /ḥ/: in this case the problem of filling a slot does not arise, since there is always a surface segment in the place of the historical /ḥ/, namely, [x]. Historical /ḥ/ could thus be postulated as underlying /x/ and the pattern correspondences could still be maintained.

Thus, the argument mentioned by Bolosky may motivate the postulation of an underlying /ʔ/ as a 'slot filler': it cannot however be cited as evidence for postulating phonemic contrast between the members of the historical guttural group (i.e. /ʔ/ : /^{c}/) or for an underlying /ḥ/.

[4]Another possible argument against abstractness involves rule ordering. If it can be established independently that the application of rule A which must refer to the surface properties of segment X precedes in the derivation rule B, which treats the same segment according to its presumed underlying features, this would obviously constitute strong evidence against postulating as an abstract segment. In many cases, however, abstract analyses recapitulate the historical developments: if the rule which has led to the neutralization of the historical underlying contrasts was the last to enter the language, and if rule reordering has not taken place, there will be no data which will enable us to construct this type of an argument. Furthermore, such argument cannot be invoked if the relevant rules cannot be proven independently to be crucially ordered one with respect to the other. This is, for example, the case with the a-Insertion Rule, which is presumably triggered by /ḥ/, and the Vowel Epenthesis Rule (i.e. the rule inserting a vowel to break the impermissable sequences $\#C\left[\begin{smallmatrix}C\\ +low\end{smallmatrix}\right]\left[\begin{smallmatrix}C\\ +low\end{smallmatrix}\right]C$, which no longer refers to historical /ḥ/ as a [+low] consonant).

Notice, however, that an argument based on rule ordering can be used only negatively, namely to demonstrate that a certain abstract analysis cannot be maintained. The nonavailability of such argument does not, of course, constitute positive evidence in favor of an abstract analysis.

[5]Since the number of morphemes where the a surfaces is approximately equivalent to those which do not undergo the rule, it seems preferable to refer to the situation as 'ad hoc subcategorization' rather than a 'minor rule' (see Kenstowicz and Kisseberth 1977, p. 126). Essentially, both involve arbitrary marking of morphemes; the only difference is that in the case of minor rules one class can be identified as regular and unmarked.

[6]Instead of the subcategorization analysis the CaCa(x) - CoCea(x) alternation could be represented by suppletion; i.e. by listing both forms in the lexicon. In both analyses memorization is involved: one has either to memorize the fact that a certain morpheme is marked with a diacritic (a fact which is totally unpredictable) or two allomorphs of a certain morpheme have to be memorized (in a suppletive analysis).

The analysis involving subcategorization seems preferable for the following reasons:

a) Whereas a suppletive alternation (as defined traditionally) would hardly be expected to be productive (e.g. extend to new items) the alternation of the type CaCa(x) - CoCea(x) is productive at least to a certain degree.

b) Two aspects of predictability should be distinguished. One has to do with whether the application of the rule is predictable, whereas the second involves the pattern of alternation. In cases of ad hoc subcategorization what is unpredictable is whether a certain morpheme will manifest the alternation or not (i.e. whether the rule will apply). Once, however, the rule applies the pattern of alternation is totally predictable: if a morpheme of the type CaCa is marked we 'know' that the alternant will be of the configuration CoCea. From this second aspect it is as predictable as a major rule. For the traditional cases of suppletion (is-are, go-went), both their occurrence as well as their form (pattern of alternation) are totally unpredictable. The cases of the type CaCa(x) - CoCea(x) seem to be somewhere in between suppletive paradigms, on one hand, and totally predictable alternations on the other. They share properties with both. Since they seem, however, to differ essentially from clear cases of suppletion (in the sense that they are both productive and predictable at least to some extent), I prefer the ad hoc subcategorization to the suppletive analysis.

For the purposes of the present argument the question as to which of the two approaches is more adequate is not of crucial importance; both imply that for the native speaker of MH the alternations discussed involve arbitrariness, as opposed to the abstract analysis, which claim that they are totally predictable.

[7]The various verb patterns of Hebrew (i.e. $C_1iC_2eC_3$, $hiC_1C_2iC_3$, etc.) are generally associated with certain semantic properties such as causativity, incohativeness, etc. The variability found in subjects' responses (licloax lecaleax, limsox lemasex) is, however, irrelevant for the purposes of the present issue; with respect to the application of the a-Insertion rule all subjects manifested a unanimous behavior.

[8]For example, the historical distinction between infinitives of the type $la^{C}aCoC$: $le?eCoC$, which depended on the contrast /?/ : /^{C}/ is not maintained even though it is reflected in the orthography.

REFERENCES

BAR ADON, A. (1959). Children's language in Israel. Unpublished Ph.D. thesis, University of Jerusalem (in Hebrew).

BARKAI, Malachi (1972). Problems in the phonology of Israeli Hebrew. Unpublished Ph.D. thesis, University of Illinois, Urbana.

BERKO, Joan (1958). The child's learning of English morphology. Word 14 150-177.

BOLOZKY, Shmuel (1972). Categorical Limitations on Rules in the Phonology of Modern Hebrew. Unpublished Ph.D. thesis, University of Illinois, Urbana.

BRAINE, Martin D. S. (1974). On what might constitute a learnable phonology. Language 50 270-299.

BRAME, Michael (1972). On the abstractness of phonology: Maltese ^{c}. In M. Brame (ed.). Contributions to generative phonology 22-61. University of Texas press.

CROTHERS, J. (1971). On the abstractness controversy (POLA reports, 2nd series 12) Berkeley, Phonology Laboratory, Department of Linguistics.

HOOPER, Joan (1976). An introduction to natural generative phonology. New York; Academic Press.

KIPARSKY, Paul (1968). How abstract is phonology? Bloomington: Indiana University Linguistic Club.

KIPARSKY, Paul (1971). Historical linguistics. In W. O. Dingwall (ed.). A survey of Linguistic Science 576-649. College Park: University of Maryland.

KIPARSKY, Paul (1975). What are phonological theories about? In David Cohen and Jessica Wirth (eds.). Testing Linguistic Hypotheses, 187-209. New York: Wiley.

KISSEBERTH, C. W. (1969). On the abstractness of phonology: an evidence from Yawelmani, Papers in Linguistics, 1 248-282.

OHALA, John (1973). On the design of phonological experiments. Monthly Internal Memorandum 25-62. Berkeley: Phonology Laboratory, Department of Linguistics.

ORNAN, Uzzi (1972). Ordered rules and the so called phonologization of ancient allophones in Israeli Hebrew. In Luigi Heilmann (ed.) Proceedings of the eleventh international congress of linguists. Bologna-Florence.

APPENDIX A

a) To wear a sock is liGRoV GeReV.
To wear a belt is

R _____ (laXGoR XaGoRa)

To put on a tie is

R _____ (la?aNoV ?aNiVa)

What would be similar constructions to convey the following ideas:

He wears a hat. ('a hat' is kova)

R _____

He puts on a blanket. ('a blanket' is smixa)

R _____

b) To sculp a statue is leFaSeL PeSeL
To draw a picture is

R _____ (leCaYeR CiYuR)

How would you say:
He makes a mask. ('a mask' is masexa)

R _____

He makes a plate. ('a plate' is calaxat)

R _____

c) To take off the peel is leKaLeF KLiPa.
How would you say:

He removes the seeds. ('a seed' is garin)

R _____

d) How would you say 'to put in the middle'. ('middle' is emca)

R _____

e) How would you say 'to make glass'. ('glass' is zxuxit)

R _____

Note: For readers who are not familiar with Hebrew, I have indicated in brackets the relevant Hebrew words.

Studies in the Linguistic Sciences
Volume 8, Number 1, Spring 1978

HEADLESS RELATIVE CLAUSES IN QUECHUA

Peter Cole and Wayne Harbert and Gabriella Hermon

A wide variety of hypotheses regarding the underlying structure of relative clauses have been put forward in the generative literature. Among the more imaginative-and, at first glance, less plausible proposals-is that of Michael Brame, who suggested that restrictive relative clauses (with surface heads) are underlyingly headless (i.e., have dummy heads). Surface heads are formed by a rule which copies the relativized NP from the embedded clause into the matrix clause.

Brame's proposal gains considerable plausibility in the light of recent work on relativization in Quechua. The purpose of this paper is to demonstrate the existence of surface headless relative clauses in three members of the Quechua language family: Ancash, Huanca and Imbabura. Evidence for headless structure is based on case marking, word order and the distribution of relative clause types. The results derive primarily from field work by the authors. The existence of surface headless relative clauses in Quechua has not been reported previously in the literature.

The appearance in surface structure of just that structure proposed by Brame as underlying restrictive relative clauses generally constitutes significant support for Brame's proposal. It suggests that Brame's analysis of English deserves new and serious consideration.

The purpose of this paper is to argue for the existence of "headless" relative clauses in three dialects of Quechua: Imbabura (a Quechua II dialect spoken in northern Ecuador), Ancash and Huanca (both Quechua I dialects spoken in central Peru).[1] The existence of such relative clauses in Quechua is of considerable theoretical importance because it lends support to a recent claim that relative clauses are headless underlyingly even in languages like English where heads invariably occur in surface structure.

We shall begin by examining briefly the motivation for positing underlyingly headless relative clauses in English. We shall then turn to Quechua and show that there is compelling evidence for the existence of surface headless relative clauses in each of the above-mentioned dialects.[2] The clear existence of headless relative clauses in Quechua provides important support for the proposal that relative clauses are underlyingly headless cross-linguistically.[3]

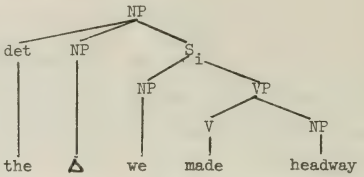
Headless Relative Clauses in English

Recent studies of relativization in English suggest that English relative clauses are headless in underlying structure.[4] According to

this hypothesis the underlying structure of (1) would be roughly (2), (in which there is a dummy head[5]):

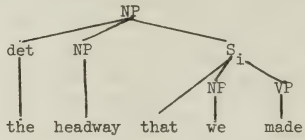
(1) the headway that we made

(2)



A rule of promotion copies the object NP into head position, yielding (3):

(3)



The evidence for such a derivation in English is quite convincing.[6] Heads of relative clauses behave syntactically (with respect to a variety of rules and constraints) as though they were constituents of the embedded clause (S_i) rather than the matrix clause. Such behavior is expected if they originate in the embedded clause and are later promoted to constituency in the matrix clause. But it is unexpected and would constitute an irregularity in the grammar if heads were matrix clause constituents in (both surface and) underlying structure.[7]

Let us examine an example of a head behaving as though it were an original constituent of (S_i). Consider relative clauses on the head noun headway, which in environments other than relative clause constructions can appear only as the object of make, as is illustrated in (4) and (5):

(4) We made satisfactory headway

(5) *The headway was satisfactory

When headway is modified by a relative clause, however, the restrictions on its distribution are waived. Compare the ungrammatical (5) with the grammatical (6):

(6) The headway [which we made] was satisfactory

The grammaticality of (6) is not predicted if we assume that (1) derives from an underlying structure in which the head originates outside the embedded S_i. But if (1) derives from a structure similar to (2), the grammaticality of (6) is predictable. Headway is grammatical as a head

NP in (1) according to that analysis because it originates as the object of make. Schachter provides a variety of arguments of similar form showing that relative heads with pronouns, picture noun reflexives and reciprocals all behave as though they originate within the relative clauses rather than as heads.

If relative clauses with heads derive universally from headless structures (i.e., structures with dummy heads), it would seem likely that there would be languages in which relative clauses are headless in surface structure.[8] We shall now turn to evidence that various dialects of Quechua provide instances of such languages.

Imbabura[9]

We shall first consider relativization in Imbabura. Before presenting the arguments for headless relative clauses in Imbabura, it will be helpful to review some general facts about relativization in this dialect. In Imbabura, as in other varieties of the Quechua language family, relative clauses, and subordinate clauses generally, appear in nominalized form. The choice of nominalizer in Imbabura is determined by the temporal relationship of the relative clause and the matrix clause. This is illustrated in (7)-(9):

(7) Relative Clause with Present Nominalizer

[$\emptyset_i$ puñu-ju-j] wawa$_i$ mana cai-pi-chu
[sleep-prog-present nom] child not this-in-neg
'The child who is sleeping is not here.'

(8) Relative Clause with Past Nominalizer

[$\emptyset_i$ puñu-shca] wawa$_i$ mana cai-pi-chu
[sleep-past nom] child not this-in-neg
'The child who was sleeping is not here.'

(9) Relative Clause with Future Nominalizer

[$\emptyset_i$ puñu-na] wawa$_i$ mana cai-pi-chu
[sleep-future nom] child not this-in-neg
'The child who will sleep is not here.'

In contrast to other dialects, when a head is present, the choice of nominalizer does not depend on the grammatical role (subject, direct object, etc.) of the relativized noun phrase.[10] For instance, -shca may be used in the relativization of transitive subjects and -j may be used in the relativization of direct objects. This is illustrated by (10) and (11):

(10) Relativization of Transitive Subject with -shca

[$\emptyset_i$ warmi-ta juya-shca] runa$_i$ aicha-ta micu-ju-n
[woman-acc love-past nom] man meat-acc eat-prog-3
'The man who loved the woman is eating meat.'

(11) Relativization of Direct Object with -j

[runa $\emptyset_i$ juya-j] warmi$_i$ aicha-ta micu-ju-n
[man love-present nom] woman^i meat-acc eat-prog-3
'The woman whom the man loves is eating meat.'

As would be expected in an SOV language, in relative clauses where a head is clearly present, it appears to the right of the embedded clause. this is represented schematically in (12):

(12) $NP[\ S_i[\ ]S_i\ NP[\ ]NP\ ]NP$

It should be noted, furthermore, that word order in relative clauses, in contrast with matrix clauses, is strictly SOV. Compare (13a) and (13b) with (14a) and (14b):

(13) Word Order in Relative Clauses

(a) SOV Order

[$\emptyset_i$ aicha-ta micu-j] warmi$_i$
[meat-acc eat-present nom] woman

*(b) SVO Order

[$\emptyset_i$ micu-j aicha-ta] warmi$_i$
[eat-present nom meat-acc] woman
'A woman who eats meat'

(14) Word Order in Matrix Clauses

(a) SOV Order

warmi-ca aicha-ta micu-n
woman-topic meat-acc eat-3
'The woman eats meat.'

(b) SVO Order

warmi-ca micu-n aicha-ta
woman-topic eat-3 meat-acc
'The woman eats meat.'

Case marking inside relative clauses follows the same principles that operate in main clauses. Subjects receive zero case marking, direct objects are marked with the accusative case marker -ta (or its allomorph -da after nasals). Case marking applies within the relative clause on the basis of grammatical relations within that clause. In addition, the relative clause as a whole receives case marking on the basis of its grammatical role in the matrix clause. These processes are illustrated in (15):

(15) parla-rca-ni [$\emptyset_i$ runa-ta ricu-shca] warmi-wan
speak-past-1sg [man-acc see-past nom] woman-with
'I spoke with the woman who saw the man.'

In (15) runa 'man' receives accusative case marking because it is the direct object within the embedded clause. The relative clause as a whole is marked with the comitative postposition -wan, the appearance of which is governed by the matrix verb parla 'speak'. Case marking within the relative clause differs from matrix case marking in only one way. The accusative case marker -ta is optional inside relative clauses though it is obligatory in matrix clauses, as shown by (16a), (16b) and (17):

(16) -ta Suppression in Relative Clauses

(a) Unsuppressed -ta

[$\emptyset_i$ wagra-ta michi-j] wawa$_i$
[cow-acc herd-pres nom] child
'child who herds cows'

(b) Suppressed -ta

[$\emptyset_i$ wagra michi-j] wawa$_i$
[cow herd-pres nom] child
'child who herds cows'

(17) Absence of -ta Suppression in Matrix Clauses

wawa-ca	*wagra wagra-ta	michi-n
child-topic	*cow-Ø cow-acc	herd-3

'The child herds cows.'

We shall now turn to headless relative clauses. The clearest and most straightforward examples of headless relative clauses in Imbabura are in sentences like those of (18):

(18) Headless Relativization of Embedded Direct Objects

(a) [wambra wagra-ta randi-shca] wañu-rca
[boy cow-acc buy-past nom] die-3 past
'The cow which the boy bought died.'

(b) [runa alcu-ta jatu-shca] ali-mi
[man dog-acc sell-past nom] good-validator
'The dog which the man sold is good.'

There are two clear cut arguments that these relative clauses are in fact headless.[11] Both arguments are based on a fundamental structural difference between relative clauses with heads and those without heads. Cross-linguistically, in those relative clauses with heads, the head is, in derived structure at least, a constituent of the matrix

rather than the embedded clause. It appears either immediately to the left or immediately to the right of the embedded clause, the position typically correlating with basic word order in the language in question (to the left of the embedded clause in VO languages and to the right in OV languages). In contrast, in headless relative clauses, unless rules changing its position apply, the relativized NP appears in its normal position within the relative clause. It is treated by case marking and other rules coding grammatical role as a constituent of the embedded rather than matrix clause.

The first argument that the relative clause constructions in (18) are headless has to do with word order. The underlined NPs in (18) appear in the middle of the embedded clause rather than to the left or right, as would be expected if they were heads. Compare (18) with (19):

(19) Relativization of Embedded Direct Object (with Head)

(a) [wambra $\emptyset_i$ randi-shca] wagra$_i$ wañu-rca
[boy buy-past nom] cow die-3 past
'The cow the boy bought died.'

(b) [runa $\emptyset_i$ jatu-shca] alcu$_i$ ali-mi
[man sell-past nom] dog good-validator
'The dog the man sold is good.'

In (18) the relativized NP appears in the normal position for a direct object within a relative clause, that is to say, between the subject and verb. In contrast, in (19) the underlined NP appears to the right of the embedded verb. The differences in order between (18) and (19) are unexceptional if the sentences of (18) involve relative clauses without heads and those of (19) with heads.

The second argument is based on case marking. In (18) the underlined NP receives accusative case marking. This would be expected on the basis of the case marking process described previously if that noun phrase is a constituent of the embedded clause, but would be unexpected if it were a head, which is a main clause constituent.

Thus, we conclude that the sentences of (18) illustrate headless relative clauses.

Having established the existence of headless relative clauses in Imbabura, we would now like to examine briefly the distribution of the construction. In the preceding examples we showed the headlessness of certain instances of relativization where the relativized NP is a direct object within the embedded clause and the past nominalizer -shca is employed. This type of relative clause is represented schematically in (20):

(20) $_{NP}[\ _{S}[$ Subject OBJECT Verb-shca $]_{S}\ ]_{NP}$

It should be noted that the sentences of (18) are unambigous with regard to whether the subject or the object has been relativized.

These sentences cannot be interpreted as involving relativization of the subject NP. That is, (18a) cannot mean 'The boy who bought the cow died.'. Nor can (18b) mean, 'The man who sold the dog is good.'. Headless relativization with the nominalizer -shca is restricted to the relativization of direct objects.[12]

Thus, in headless relative clauses, unlike those with heads, the nominalizing suffix indicates both tense and the grammatical role of the relativized NP. Such a mixture of functions is not unusual in Quechua generally, but as we showed previously, it is not found in Imbabura in relative clauses with heads. We shall employ the difference in function of the nominalizing suffixes in relative clauses with and without heads to help us determine whether certain other relative clauses have or do not have heads. The difference in function of the nominalizer should be particularly helpful in examples where case marking and word order fail to distinguish whether a relative clause is or is not headless.

Let us consider the examples in (21):

(21) Headless Relativization of Embedded Subjects

(a) [wambra wagra-ta randi-j] wañu-nga caya
[boy cow-acc buy-pres nom] die-future 3 tomorrow
'The boy who is buying the cow will die tomorrow.'

(b) [runa alcu-ta jatu-j] ali-mi
[man dog-acc sell-pres nom] good-validator
'The man who is selling the dog is good.'

The sentences of (21) are similar to those of (18) in that they must be interpreted as involving the relativization of a single grammatical role, subject, in this case. Thus, (21a) cannot constitute a statement that the cow will die tomorrow nor (21b) that the dog is good. The sentences of (21) differ from those of (18) in that in (18) the relativized NP is the embedded direct object while in (21) it is the embedded subject. We shall argue that these sentences, like those of (18), are instances of headless relative clauses (as indicated by the bracketing and the caption given in the examples). Alternatively, one might want to argue that these sentences should be interpreted not as instances of headless relative clauses, but rather as instances of right branching relative clauses with heads on the left. The two possible structures are illustrated in (22):

(22) a. Headless Relative Clause

$NP[\ S[\ \underline{SUBJECT}\ Object\ Verb\text{-}j]_S\ ]_{NP}$

b. Right Branching Relative Clause

$NP[\ NP[\ \underline{HEAD}\ ]_{NP}\ S[SUBJECT\ Object\ Verb\text{-}j]_S\ ]_{NP}$

We shall argue that the sentences of (21) are not in fact right branching, but rather are headless.

The two hypotheses in question make the same predictions regarding case marking and word order. Both predict that the underlined noun phrases of (21) would receive zero case marking and appear in initial position. Thus, word order and case marking cannot be used as tests for headless relativization in these cases.

It is, however, possible to determine empirically whether the sentences of (21) are headless or right branching. It will be remembered that in clear instances of relative clauses with (non-dummy) heads, left branching relative clauses with heads on the right, relativization was possible regardless of the grammatical role of the relativized NP within S_i and regardless of the nominalizing suffix employed. If the relative clauses of (21) are right branching, it would, therefore, be expected that sentences similar to (21) but involving the relativization of the object of S_i should be grammatical. But, as the sentences of (23) show, such relative clauses are not well-formed:

(23) a. ?? wagra$_i$ [wambra $\emptyset_i$ randi-j] wañu-rca
cow [boy buy-pres nom] die-past 3
'The cow which the boy was buying dies.'

b. ?? alcu$_i$ [runa $\emptyset_i$ jatu-j] ali-mi
dog [man sell-pres nom] good-validator
'The dog the man is selling is good.'

Right branching relative clauses in which the object of the lower clause is relativized and the present nominalizer -j is employed are ungrammatical.

It might be proposed that the sentences of (23) do not constitute conclusive evidence that the right branching hypothesis is false. Perhaps in right branching relative clauses the nominalizers encode both tense and the grammatical role of the relativized NP. Relative nominalizers have both of these functions generally in most dialects of Peruvian Quechua. They also have both functions in Imbabura headless relatives, as demonstrated by the lack of ambiguity in the examples in (18). Thus, it would not be surprising if the -j nominalizer were restricted to relativized subjects and -shca to relativized objects in right branching relative clauses. Hence, according to the revised right branching hypothesis, the ungrammaticality of (23) would be due to the choice of -j as nominalizer. This proposal may be easily tested by substituting -shca and/or -na for -j in the examples in (23):

(24) a. ?? wagra$_i$ [wambra $\emptyset_i$ randi $\left\{\begin{matrix}\text{-na}\\\text{-shca}\end{matrix}\right\}$] wañu-rca
cow [boy buy $\left\{\begin{matrix}\text{-fut nom}\\\text{-past nom}\end{matrix}\right\}$] die-past 3
'The cow which the boy $\left\{\begin{matrix}\text{would buy}\\\text{had bought}\end{matrix}\right\}$ died.'

b. ?? alcu$_i$ [runa $\emptyset_i$ jatu $\left\{\begin{matrix}\text{-na}\\\text{-shca}\end{matrix}\right\}$] ali-mi
dog [man sell $\left\{\begin{matrix}\text{-fut nom}\\\text{-past nom}\end{matrix}\right\}$] good-validator
'The dog which the man $\left\{\begin{matrix}\text{will buy}\\\text{bought}\end{matrix}\right\}$ is good.'

The ungrammatical examples of (24), in which -shca and -na have been so substituted show, we believe, that the ungrammaticality of (23) is, in fact, not due to the choice of nominalizer. If the right branching hypothesis were to be maintained in the face of such facts, it would be necessary to restrict right branching relative clauses to the relativization of embedded subjects. Furthermore, headless relative clauses would have to be restricted to object relativization. Thus, the right branching analysis is fraught with irregularities and ad hoc restrictions.

But if the sentences of (21) are analyzed as headless, such restrictions are unnecessary. In headless subject relatives. the nominalizer -j is employed while in headless object relatives the nominalizer -shca is employed. The absence of sentences like those of (23) and (24) is predicted if Imbabura does not allow right branching relative clauses at all.

Thus, we conclude that in Imbabura there are two types of relative clauses, left branching and headless. Headless relative clauses may be formed on either embedded subjects or embedded objects. The choice of relativized NP is indicated by the nominalizer. A similar, though not identical pattern will be seen in Huanca, to which we shall now turn.

Huanca

The discussion of Huanca is based entirely on the excellent reference grammar for the dialect written by R. Cerrón Palomino.[13] Cerrón-Palomino describes Huanca as having both left branching relative clauses (like (25)), and right branching relative clauses.

(25) [$\emptyset_i$ lisqi-sha-yki] nuna-kak-man$_i$ kuti-y
[know-nom-2] man-the-to return-imperative
'Return to the person that you know.'

Right branching is claimed to be predominant. He notes that if this description is correct, the situation in Huanca is anomalous since there does not appear to be any other dialect of Quechua in which a similar preference for right branching relativization is to be found. We shall propose in this section that relativization in Huanca is not anomalous, but rather that it is quite similar to that found in Imbabura. We suggest that all those instances of relative clauses analyzed by Cerrón-Palomino as right branching may in fact be headless. Thus, the relativization strategies in Huanca are reanalyzed as headless and left branching, the same two strageties found in Imbabura. The preference for putative "right branching" over left branching is treated in our analysis as a preference for headlessness over left branching. Given the widespread occurrence of headless relative clauses in Quechua generally, the latter preference would not seem particularly surprising.

We would like to begin by arguing that certain relative clauses analyzed by Cerrón-Palomino as right branching must be headless. The argument is based on word order, as was the first and most straightforward argument from Imbabura. Consider relativization of embedded direct objects when the -sha nominalizer (cognate to Imbabura -shca) is employed:

(26) Relativization of Embedded Direct Object with -sha

luwis-pa kawallu lanti-sha-n-kak alfak-ta miku-yka-n
Luis-genitive horse buy-nom-3-the alfalfa-acc eat-prog-3
'The horse that Luis bought is eating alfalfa.'

Cerrón-Palomino treats sentences like (26) as right branching. But this clearly cannot be correct because the putative head, kawallu appears in normal object position within the embedded clause. (Kawallu fails to receive accusative case marking in Huanca because the suppression of the accusative case marker is obligatory in relative clauses.) Thus, sentences like (26) must have a derived structure like (27) rather than one like (28):

(27) $_{NP}[\ _{S}[\ ]_{S}\ ]_{NP}$

(28) $_{NP}[\ _{NP}[\ ]_{NP}\ _{S}[\ ]_{S}\ ]_{NP}$

The fact that some putative right branching relative clauses must be analyzed as headless raises the question of whether all putative right branching might in fact be headless. An examination of the range of data adduced by Cerrón-Palomino provides no counterexamples to the headless hypothesis:

(29) Relativization of Embedded Subject with -sha

walash libra lanti-sha-n-kak wañu-ku-n
boy book buy-nom-3-the die-refl-3
'The boy who bought the book died.'

(30) Relativization of Embedded Subject with -q[14]

allqu chuqllu suwa-q-kak yana-m
dog corn steal-nom-the black-validator
'The dog that steals the corn is black.'

In (29) and (30) the underlined NPs, walash and allqu, might, on the basis of the data available, be analyzed either as heads or as the subjects of the embedded clauses. Ultimately, this question must be resolved by finding dialect specific tests for the constituency of walash and allqu, as we in fact did in the previous section for Imbabura. We are not yet in a position to do this for Huanca, so any conclusions drawn at this time must be tentative. It would seem likely, however, that Huanca relativization is not as anomalous as Cerrón-Palomino believed, and that headless relativization is central to relative clause formation in Huanca.[15]

Ancash[16]

Relativization in Ancash is considerably more complex than in Imbabura and Huanca. As in other dialects, left branching relative clauses are found, as in example (31):

(31) poncho-ta rura-q warmi šumaq-mi
poncho-acc make-nom woman good-validator
'The woman who makes ponchos is good.'

But, in addition, there is found in this dialect a range of case marking facts which, so far, has eluded complete analysis on our part. We are unaware of any satisfactory general account of the complexities we have encountered. In particular, we are unable to specify which grammatical processes interact to derive apparent right branching relative clauses like (32):

(32) kuya-a-mi taqay warmi-ta qotsu-q-ta
love-1-validator that woman-acc sing-nom-acc
'I love the woman who is singing.'

in which the apparent head and the nominalized verb both receive matrix case marking, and apparent right branching relative clauses like (33):

(33) noqa kuya-a-mi warmi-ta qam dansa-šqa-iki-wan
I love-1-validator woman-acc you dance-nom-2-obl
'I love the woman you dance with.'

in which the nominalized verb receives the case marking of the relativized NP in the embedded clause and the apparent head receives matrix case marking. In other examples, a noun phrase bearing embedded clause marking appears to the left of the embedded clause and the nominalized verb receives the same case marker as that noun phrase. This is illustrated by example (34):

(34) noqa urya-a nuna-ta qam qellay-ta qušqa-iki-ta
I work-1 man-acc you money-acc give-nom-2-acc
'I work with the man you give money to.'

In still further examples, the NP bears embedded case marking and the nominalized verb matrix case marking as in (35):

(35) kuya-a-mi taqay siudad-čaw noqa yača-q-ta
love-1-validator that city-loc I live-nom-acc
'I love the city I live in.'

Furthermore, the range of possibilities accepted often appears to vary idiosyncratically from verb to verb. Until the regularity behind this complex of surface forms is revealed, it will be risky to make claims about the overall structure of the relative clause in Ancash.[17]

There do, however, seem to be certain relative clauses that appear almost certainly to be headless. As in Imbabura and Huanca, the clear cases involve the relativization of direct objects:

(36) a. noqa dansa-rqa-a mario warmi-ta kuya-šqa-n-wan
I dance-past-1 Mario woman-acc love-nom-3-obl
'I dance with the woman Mario loves.'

(36) b. mario warmi-ta kuya-šqa-n wañu-šqa
Mario woman-acc love-nom-3 died-past
'The woman Mario loves died.'

The sentences of (36) appear to be completely analogous to instances of headless relativization of direct objects in Imbabura. As in both Imbabura and Huanca, the position of the underlined NPs is the one appropriate if those NPs are constituents of the embedded clause, but not if they are heads. Similarly, the accusative case marking of the italicized NPs is clearly due to their grammatical role within the embedded clause.[18] This case marking would be unexplained if these NPs were heads and, hence, matrix clause constituents. Additional examples, in which oblique NPs within the embedded clause are relativized appear in (37):

(37) a. noqa kuya-a-mi qam warmi-wan dansa-šqa-iki-ta
I love-1-validator you woman-obl dance-nom-2-acc
'I love the woman you danced with.'

b. mario wayi-čaw yačа-šqa-n rupa-šqa
Mario house-loc live-nom burn-past
'The house Mario lived in burnt down.'

The argument that relative clauses like those in (36) and (37) are headless is further strengthened by the fact that when a direct object or oblique relativized noun phrase remains in its normal position between the subject and the verb it may not receive matrix clause case marking. This is shown in the examples in (38):

(38) a. *noqa dansa-rqa-a mario warmi-wan kuya-šqa-n-wan
I dance-past-1 Mario woman-obl love-nom-3-obl
('I danced with the woman Mario loved.')

b. *kuya-a-mi qam warmi-ta dansa-šqa-iki-ta
love-1-validator you woman-acc dance-nom-2-acc
('I love the woman you dance with.')

The ungrammaticality of the sentences of (38) is particularly striking because matrix case marking is possible when the apparently analogous noun phrase appears to the left of the subject, as in (33) above. Thus, the underlined NPs in (36) do not behave as heads with respect to either word order or case marking.

Although it is possible to establish that some relative clauses are headless in Ancash, it is not possible on the basis of what we know to determine the distribution of the construction. Are headless relative clauses limited to certain instances of the relativization of objects as in (36) and (37), or is the construction pervasive in this dialect, as it is to some extent in Imbabura and to perhaps an even greater extent in Huanca?

Conclusions

We have taken pains to demonstrate that three Quechua dialects have headless relative clauses. (Our informant for a third dialect, Cochabamba, appeared to lack the construction. Headless relative clauses are apparently not found in southern Quechua.) It is of some importance that Imbabura on the one hand and the two Quechua I dialects on the other are about as distantly related genetically and geographically as any two dialects of Quechua. It would seem unlikely that either dialect acquired the construction by borrowing and, given the rareness of headless relative clauses among the languages of the world, it is improbable that the three dialects developed the construction independently.[19] Thus, it would seem that headless relativization was likely to have been found in Common Quechua, as P. Muysken has suggested. Headless relativization appears to be a relativization strategy of some importance among Quechua dialects.

We have observed that the existence of this strategy provides strong support for the Brame/Schachter hypothesis, according to which even relative clauses with heads derive from underlyingly headless constructions. In addition to the theoretical significance of headless relatives for the study of the structure of relative clauses generally, they provide a testing ground for a variety of theoretical claims. Was Ross[20] correct in defining the Complex NP Constraint on structures of the form NP-S, or should the constraint prohibit extractions out of relative clauses as such? Headless relative clauses may provide the relevant test case.

Were Keenan and Comrie[21] right in holding that the more a relative clause preserves of normal sentential form, the more it will be able to relativize "difficult" positions on the NP Accessibility Hierarchy? Are Quechua headless relatives in fact used for such purposes? These are indicative of some of the questions that Quechua headless relatives may assist us in answering.

NOTES and REFERENCES

[1]The Quechua language (family) is spoken by six to eight million persons in western South America. There are two major dialect groups, termed Quechua I and II by Alfredo Torero ("Los dialectos quechuas". Anales Cientificos de la Universidad Agraria, Vol II, No.4. (Lima, 1964), pp. 446-478.). For information on the classification of Quechua dialects, see Torero, op. cit. and Gary J. Parker ("La Clasificación Genética de los Dialectos Quechuas", Traducido por Gabriel Escobar. In: Revista del Museo Nacional. Tomo XXXII. (Lima, 1963), pp. 241-252).

The present study is based on informant work carried out in Ecuador, Peru and at the University of Illinois (Urbana). An earlier version of this paper was presented at the Workshop on Andean Linguistics held at the University of Illinois in July, 1978. The authors gratefully acknowledge the support of the National Science Foundation (grant No. SOC 75-00244 and grant No. BNS 77-27159), the Research Board and the Center for International Comparative Studies of the University of Illinois for their support of this research.

[2]Paul Schachter ("Focus and relativization". In: Language, Vol.49, No.1 (1973), pp. 19-46) cites data from Charles S. Bird (Relative clauses in Bambara". In: Journal of West African Languages, Vol.5 (1968), pp. 35-47) which strongly suggest that Bambara relative clauses are headless in surface structure. Avery D. Andrews ("A typologcal survey of relative clauses". (1972) Unpublished Manuscript) also argues for the existence of headless relative clauses in several languages. Pieter C. Muysken ("Relative clause formation in Ecuadorian Quechua". (1976) Unpublished Manuscript) makes claims similar to ours for Quechua.

[3]This is not to suggest that evidence of underlying headlessness is not necessary for English and other languages.

[4]For example, Schachter, (op. cit.), and Brame, (Ms., "A new analysis of the relative clause: evidence for an interpretative theory")

[5]As a matter of convenience, we adopt Schachter's dummy head proposal. An alternative analysis is one in which there is no head at all underlyingly. A third possibility is that the head is deleted under coreference with the relativized noun phrase. We know of no empirical consequences for the hypothesis of a dummy head rather than of underlying headlessness or head deletion. Head deletion would not, however, be compatible with Brame's data from English.

[6]These arguments are taken from Schachter (op. cit.), but are apparently due ultimately to Brame.

[7]For an alternative explanation of the English data, see Pauline Jacobson ("Some aspects of movement and deletion". In: Berkeley Linguistic Society, Vol. III (1977).

[8]It is immaterial for our pruposes whether superficially "headless" relative clauses contain a dummy head like that posited in deep structure, or whether the head is deleted prior to surface structure. For the sake of simplicity in representation, we will assume that the dummy head is deleted.

[9]Data in this section is based on informant work conducted by Peter Cole and Janice L. Jake. We would like to thank our principal informant, Carmen Chuquín, for her generous assistance. The data presented here represents the Quechua spoken in Mariano Acosta, in the eastern part of the province of Imbabura. We do not know whether all of our claims are true for Imbabura generally. In particular, the uses of the nominalizing suffixes reported here differ from those discussed by Muysken (op. cit.). Possibly the informants Muysken consulted were speakers of a subdialect of Imbabura with rules different from those in Mariano Acosta.

[10]In most Quechua dialects both grammatical role and tense are involved in the choice of nominalizers. See Muysken (op. cit.) for a

survey of nominalizer choice in a variety of dialects. Typically, the nominalizer cognate to Imbabura -j is used with subjects and -shca with direct objects. A similar pattern is found in Imbabura, but only in surface headless relative clauses. This is discussed further below.

[11]That is to say, they have dummy heads or no heads. See Fn. 8.

[12]And, possibly, intransitive subjects. We have elicited such sentences as the following:

(i) jari shamu-shca yarjachi-n
man came-past nom hungry-3
'The man who came is hungry.'

[13]Rudolfo Cerrón-Palomino ("Gramatica Quechua Junin-Huanca", Ministerio de Educación/Instituto de Estudios Peruanos (Peru, 1976)).

[14]The -q nominalizer is cognate to Imbabura -j.

[15]-na relatives might be useful in testing our claims, but unfortunately, all of Cerrón-Palomino's examples involve pronominal subjects (so-called transitions). Thus, a word order test cannot be used:

(i) ponchu trula-ku -na -n -kaq -ta apamu-y
poncho put on-refl -nom -3 -the -acc bring-imperative
'Bring the poncho that he is going to put on.'

Note that the sole reflex of the subject of trulaku- is the third person possessive suffix -n, which is appended to the nominalizer -na. Such examples do not tell us whether ponchu would precede or follow a subject NP.

[16]We would like to express our appreciation to Celestino Figueroa Bedón, originally from the town of Taricá, Department of Ancash for information on Ancash.

[17]According to Peter Landerman (personal communication) David Webber has found a similar case marking pattern in relative clauses in other central Peruvian varieties of Quechua. These facts are to be discussed in Webber's forthcoming Masters Thesis.

[18]Though there appears to be some variation from one sub-dialect of Ancash to another, in the sub-dialect of our informant (Huaraz-region) the choice of the nominalizing suffixes, -q and -sqa, depends solely on the grammatical role of the relativized NP. Relativized subjects (and locatives) take the -q nominalizer, whereas direct objects, indirect objects and obliques take the -sqa nominalizer.

[19]We would like to show that geographically intervening dialects lack headless relatives, but we do not have sufficient information to do so at this time.

[20]John R. Ross ("Constraints on variables in syntax". Ph.D. Thesis (1967)). MIT.

[21]Keenan, Edward L. and Bernard Comrie ("NP accessibility and universal grammar". In: Linguistic Inquiry, Vol. 8 (1972), pp. 63-100.)

REFERENCES

Andrews, Avery D. 1972. A typological survey of relative clauses. Unpublished manuscript.
Bird, Charles S. 1968. Relative clauses in Bambara. Journal of West African Languages 5.35-47.
Brame, Michael. 197 . A new analysis of the relative clause: evidence for an interpretative theory. Unpublished manuscript.
Cerrón-Palomino, Rudolfo. 1976. Gramatica Quechua Junin-Huanca. Peru.
Jacobson, Pauline. 1977. Some aspects of movement and deletion. Berkeley Linguistic Society 3.
Keenan, Edward L. and Bernard Comrie. 1972. NP accessibility and universal grammar. Lingustic Inquiry 8.63-100.
Muysken, Pieter C. 1976. Relative clause formation in ecuadorian Quechua. Unpublished manuscript.
Parker, Gary J. 1963. La Clasificación Genética de los Dialectos Quechuas. Tracucido por Gabriel Escobor. Revista del Museo Nacional 32.241-252.
Ross, John R. 1967. Constraints on variables in syntax. Ph.D. Thesis.
Schachter, Paul. 1973. Focus and relativization. Language 49.19-46.
Torero, Alfredo. 1964. Los dialectos quechuas. Anales Cientificos de la Universidad Agraria 2.446-478.

Studies in the Linguistic Sciences
Volume 8, Number 1, Spring 1978

ON THE ACQUISITION OF SUBJECTHOOD*

P. Cole W. Harbert G. Hermon S.N. Sridhar

This paper constitutes an examination of the process by means of which subject properties are acquired by a noun phrase not previously having those properties. The evidence presented indicates that behavioral subject properties (like the ability of an NP to delete under Equi) are acquired historically prior to subject coding properties (like nominative case marking and control of verb agreement).

I. Introduction

The purpose of this paper is to examine cross-linguistically the historical principles governing the acquisition of subject properties by a noun phrase.[1] Several investigators, including Keenan (1976), have shown that these properties do not always converge on a single noun phrase. In this paper we are especially concerned with the question of whether certain kinds of properties associated with subjecthood are acquired prior to other kinds of properties. In particular, are morphosyntactic (i.e. coding) properties like nominative case and control of verb agreement acquired before or after transformational (i.e. behavioral) properties like control of reflexivization and deletability by rules normally affecting only subjects? We shall argue that behavioral properties are consistently acquired prior to coding properties. Evidence in support of this claim will be presented from three groups of languages: Germanic, Polynesian, and Georgian. For each language (or group of languages), we will consider both behavioral and coding properties of subjecthood at various historical stages. Among the behavioral properties considered will be the capacity to trigger reflexivization in languages in which only subjects control that process, the capacity to undergo and to trigger rules of deletion under identity which are normally restricted to subjects, and the capacity to undergo subject raising rules. The coding properties considered include subject case marking and control of verb agreement.[2] It should be noted that we do not consider initial position among the subject coding properties. Rather, following Timberlake (1976: 560), we regard initial position as primarily indicative of topicality rather than subjecthood. While there is a frequent correlation between topicality and subjecthood, the two do not necessarily coincide. Thus, initial position will not figure in the arguments that follow.[3]

For each language (family), we shall argue that at an early stage in the development of the construction (Stage A), the noun phrase in question displays no subject properties, either behavioral or coding. At a later time (Stage B), the analogous noun phrase exhibits behavioral,

but not coding properties. If coding properties are found at any stage, they are acquired subsequent to behavioral properties (Stage C). We propose that such a sequence is the unmarked (and perhaps universal) order of acquisition of subject properties cross-linguistically. The bulk of the paper will be devoted to arguments that such a sequence is an accurate portrayal of the acquisition of subject properties in the languages discussed here. We know of no instances in which coding properties are acquired before behavioral properties.

II. Germanic Languages

We shall begin by tracing the development of two constructions in Germanic: dative experiencer constructions as exemplified in (1):

(1) Dative experiencer in Germanic (example from German)

Mir gefallen diese Damen
me-dat please these ladies
'I like these ladies.'

and passive sentences with verbs taking non-accusative objects, as in (2):

(2) Passivized objects of verbs taking dative objects in Germanic (example from German)

Uns wird von der Polizei geholfen
us-dat aux-3sg by the police help-past part
'We are helped by the police.'

We shall argue that at Stage A, no subject properties are associated with dative experiencers or passivized non-accusative objects.[4] At Stage B, these noun phrases come to exhibit the behavioral accoutrements of subjecthood, but fail to display subject coding properties. Finally, at Stage C, these noun phrases, in some Germanic languages, come to possess both syntactic and morphosyntactic subject properties.

The argument from Germanic is complicated by the number of languages involved and by the fact that the parent language, Proto-Germanic, which corresponds to Stage A, is unattested. We must produce an argument, based on a comparison of the attested Germanic languages, that at this unattested stage no subject properties were associated with dative experiencers and passive subjects.

Germanic consists of three subfamilies, East Germanic, North Germanic and West Germanic.[5] The sole extensively documented member of the East Germanic subgroup, Gothic, is represented by a relatively small corpus, the bulk of which is translated from Greek. Thus, only a very limited amount of evidence is available which bears on the status of the noun phrases in question in East Germanic. There is at least some evidence from Gothic, however, indicating that dative experiencers had no subject properties. Two rules, Equi and Conjunction Reduction, which normally applied only to subjects, treated the nominative NP in experiencer constructions, rather than the dative experiencer, as

subject. This is illustrated in (3a) and (3b):

(3) a. Nominative deletes under Equi (example from II Cor 5:9)

inuþ-þis	usdaudjam...	waila	galeikan	imma
because of-this	strive(1pl)	well	V[dat](inf)	him(dat)

'Because of this we strive to please him well.'

b. Nominative undergoes conjunction reduction (ex. from Thess 4:1)

hvaiwa	skuluþ	gaggan jah	galeikan	guda
how	should(2pl)	go(inf) and	V[dat](inf)	god(dat)

'how you should live and please God.'

As illustrated by (4), the dative experiencers in such constructions also failed to control verb agreement:

(4) Dative experiencers do not control verb agreement (ex. from I Thess 3:1)

galeikaida	uns	ei	biliþanai	weseima
V[dat](3sg)	us(dat)	that	left	might be

'It pleased us that we might be left.'

Thus, we conclude that in Gothic, the nominative, rather than the dative experiencer in experiencer constructions, had both the coding and the behavioral properties of subjects.[6] The significance of this evidence for the development of experiencer constructions in Germanic will be seen below.

In the earliest extensive North Germanic texts (Old Icelandic), behavioral subject properties appear to have been associated with the dative experiencers of only one verb, þykkia 'seem'. In an extensive study of reflexive pronouns in Old Icelandic, Rose (1976) lists several instances in which dative experiencers of this verb serve as antecedents of reflexive pronouns (a function normally restricted to subjects). This is illustrated by example (5):

(5) Dative experiencer of þykkia controls reflexivization (Njall.: 278)

honum	þóttir	þú	hafa	haft	við	sik	fjǫrráð
him(dat)	seem(3sg)	nom	aux (inf)	part	prep	refl	death-plot

'He thought you to have had a death-plot against him.'

There appear to be no attestations in which such reflexivization is controlled by passivized non-accusative objects, or by experiencer NPs other than those occurring with þykkia.[7,8] Thus, the earliest Icelandic texts appear to represent a transitional period, between Stage A and Stage B. At this stage, at least one behavioral subject property was associated with the dative experiencer of a single predicate.

In contemporary Icelandic, a substantial change appears to have taken place. Andrews (1976) notes a variety of behavioral properties associated both with a wide range of experiencers and with passivized non-accusative objects in Modern Icelandic. Example (6) illustrates

dative experiencers and (7) illustrates passivized non-accusative objects:

(6) Dative experiencers

a. tröllí-nu svelgdist á stúlku-nni (A 10c)
troll-the(dat) mis-swallowed on girl-the(dat)
'The troll swallowed the girl wrong.'

b. mér líkar/líka þeir bílar (A 10d)
me(dat) like(3sg/pl) those cars(nom)
'I like those cars.'

(7) Passivized non-accusative objects

a. mér var bjargað (A 13a)
me(dat) aux V[dat](supine)
'I was saved.'

b. mín var beðið (A 13d)
me(gen) aux V[gen](supine)
'I was waited for.'

As in Old Icelandic, the trigger for reflexivization must be a subject. Example (8) shows that dative experiencers control reflexivization:

(8) Dative experiencer controls reflexivization

henni svelgdist á stéikinni sinni (A 39b)
she(dat) mis-swallowed on steak her(refl)
'She swallowed her steak wrong.'

Andrews does not provide examples parallel to (7) involving reflexivization. He does give examples involving the passivization of indirect objects. Example (9a) illustrates an active sentence with both a direct and an indirect object. In (9b), the indirect object has been passivized and controls reflexivization.[9]

(9) a. Active bitransitive sentence

þeir seldi honum drengina (A 40a)
him(dat) sold him(dat) boys-the(acc)
'They sold him the boys.'

b. Passivized indirect object controls reflexivization

honum voru seldir drengirnir af fraendum
him(dat) aux(inf)sold(part) boys-the(nom) by relatives

sínum/hans (A 42b)
refl/anaphoric

'He$_i$ was sold the boys$_j$ by his$_i$/self's$_i$ relatives.'

Dative experiencers and passivized non-accusative objects can also be raised by subject-to-object raising, a process which normally applies only to subjects. This is shown in (10)-(11):

(10) Dative experiencer undergoes SOR

ég	tel	honum	líka	þeir	bílar	(A 19a)
I	believe	him	inf	those	cars(nom)	

'I believe him to like those cars.'

(11) Passivized non-accusative object undergoes SOR

ég	tel	þeirra	hafa	verið	bedið	(A19d)
I	believe	them(gen)	aux(inf)	aux(sup)	V[gen](supine)	

'I believe them to have been visited.'

Furthermore, experiencers and passivized dative objects may be the victims of Equi:[10]

(12) Dative experiencer deletes under Equi

tröllið	vonast	til að	svelgjast	ekki	á	
troll the(nom)	hope	comp	mis-swallowed	not	on	

stúlkinni (A 26a)
girl the

'The troll hopes not to swallow the girl wrong.'

(13) Passivized dative object deletes under Equi

ég	vonast til að verða	bjargað	(A 26d)	
I	hope	comp	aux(inf)	V[dat](supine)

'I hope to be saved.'

Thus, the data from Andrews indicate an extension in the history of Icelandic of the syntactic properties of subjecthood exhibited by experiencers and passivized non-accusative objects. It should be noted, however, that these noun phrases do not show any morphosyntactic subject properties. In experiencer constructions like (6b) (repeated):

(6) b. mér líkar/líka þeir bílar (A 10d)
me(dat) like(3sg/pl) those cars(nom)
'I like those cars.'

þeir bílar, not mér, appears in nominative case. Similarly, verb agreement is triggered by the nominative noun phrase, or fails to apply. This is illustrated by (6b) as well. Hence, Modern Icelandic appears to be an instance of Stage B. Dative experiencers and passivized non-accusative objects acquired behavioral properties typical of subjects, but not coding properties.

We will now turn to evidence for the transition from Stage B to stage C in North Germanic. First, let us consider Faroese, which is closely related to Icelandic. In Faroese, accorking to Lockwood (1964: 103),

passivized non-accusative objects in some types of passive constructions may assume nominative case. Compare (14)-(17) with Andrews' examples from Modern Icelandic:

(14) Active dative object sentence

teir fagnaðu Depilsmonnum vael
they V[dat] Depilsmen(dat) well
'They welcomed the Depilsmen well.'

(15) Passivized dative object becomes nominative obligatorily

Depilsmenn vǫ́ru vael fagnaðir
Depilsmen(nom) aux(3pl) well V[dat](part)
'The Depilsmen were well received.'

(16) Passivized dative object becomes nominative optionally

hon var boðin í brúðleyp
she(nom) aux V[dat](part) to wedding
'She was inviǿed to a wedding.'

(17) Passivized dative object remains dative

maer var ikki boðið at koma
me(dat) aux neg V[dat](supine) comp inf
'I was not invited to come.'

Examples (14)-(17) show that passive subjects of verbs taking dative objects in Faroese have begun to take on morphosyntactic subject properties of Stage C.[11]

In North Germanic languages other than Icelandic and Faroese, dative and accusative case have merged. These languages therefore no longer have non-accusative objects. However, the acquisition of subject coding properties by historically non-nominative experiencers is well advanced in them. Examples will be limited to Swedish, but similar examples from Danish and Norwegian could be supplied as well. The examples in (18) and (19) and similar ones cited by Evans (1975) demonstrate that in Old Swedish, as in Icelandic, such experiencers did not have the coding properties of subjects, although they had at least some behavioral subject properties.

(18) Dative experiencers do not control verb agreement

thz tykker them allom wara thz besta
that V[dat](3sg) them(3pl dat) all be the best
'That seemed to all of them to be the best.' (Rim-Krǫniker I p. 391)

(19) Dative experiencers control reflexivization

honom thykte sik wara j enom lystelikom stadh
him(dat) V[dat] refl be in a pleasurable city
'He thought himself to be in a pleasurable city.' (Siaelinna Thrøst, p. 280, ii. 17-18)

In modern Swedish, as in Danish and Norwegian, these historically non-nominative experiencers have acquired the subject coding properties of nominative case and control of verb agreement. A partial list of predicates which have come to occur with nominative subject-experiencers is given in (20):

(20) Historically non-nominative experiencers acquire nominative case and control of verb agreement

a.
jag drömmer	'I dream'
jag tycker	'I think'
jag kommer i hug	'I recall'
jag bör	'I ought'
jag är varm	'I am warm'
jag samtycker	'I agree'
jag saknar	'I lack'
jag längtar	'I miss'
jag behöver	'I must'

This list could be readily extented. Thus, Swedish and other contemporary Scandinavian languages are illustrative of stage C.[12]

Although in North Germanic there is no direct evidence of a stage at which no subject properties are associated with dative experiencers and passive subjects, the attested development from behavioral properties with a limited range of predicates to behavioral, and ultimately coding properties with a wide range of predicates strongly suggests that this was the case at an earlier stage in Germanic. Evidence from West Germanic, to which we now turn, supports that hypothesis.

We shall discuss data from two West Germanic languages, German and English. In German dative experiencers and passivized non-accusative objects display no subject properties, and we have found no evidence that the situation in Modern German differs from that at earlier stages in the development of the language. The absence of morphosyntactic subject properties is shown in (21)-(22):

(21) a. Active dative object sentence

Die Polizei	hilft	uns
the police	help(sg)	us(dat)

'The police help us.'

b. Passivized dative object remains dative and fails to control verb agreement

Uns	wird	von	der	Polizei	geholfen
us(dat)	aux(3sg)	by	the	police	V[dat](past part)

'We are helped by the police.'

(32) Passivized dative objects do not delete under relative clause reduction

*Der vom Lehrer geholfene Junge bekam eine gute Note
the by the teacher help(part) boy got a good grade
('The boy helped by the teacher got a good grade.')

(33) Dative experiencers do not delete under relative clause reduction

*Der das Buch gefallende Junge sitzt da in der Ecke
the the book V[dat](part) boy sits there in the corner
('The boy liking the book is sitting there in the corner.')

Thus, with respect to three behavioral properties of subjecthood, neither passivized non-accusative objects nor dative experiencers behave as subjects in German. For further discussion of the non-subjecthood of passivized dative objects, see Reis (1973), Ebert (1975) and Breckenridge (1975). We conclude that in German these noun phrases show neither coding nor behavioral subject properties. The relevance of this fact for our argument will be discussed shortly.

We will now turn to English. As was the case with Swedish, neutralization of the dative/accusative distinction in English causes us to limit our discussion to experiencers. In Old English, there are sporadic instances in which experiencer NPs behave as subjects, although they do not exhibit subject coding properties. In (34) for example, the accusative-experiencer verb langian 'to desire, long for' appears embedded under the subject-to-subject raising verb onginnen 'begin'. The raised experiencer retains its accusative case.

(34) Accusative experiencer undergoes SSR

þā ongan hine eft langian on his cȳþþe
then began him(acc) again V[acc](inf) for his kin
'Then he began to long for his kin.' (Blick. Hom. 113:15, A.D. 971)

This example substantially antedates the first appearance of langian with a nominative experiencer, for which Gaaf (1904) gives the date 1320.

As was noted by Gaaf (1904)and Harris (1975) inter alia, the experiencer NPs occurring with a number of verbs, such as liken and dremen, had subject behavioral properties in early Middle English, but had not yet acquired subject coding properties. Note the accusative case of þe 'thee' in (35a) and of him in (35b):

(35) Accusative experiencers

a. ..a foreward þat þe mai full well like
an agreement that you(acc) may full will V[acc]
'an agreement that may full well please you' (Dame Sirith: 255, A.D. 1275)

(35) b. of þat him drempte in prisun
of that him(acc) V[acc] in prison
'He dreamed of that in prison.' (Genesis and Exodus, 2116, A.D. 1250)

Although the experiencers in (35a) and (35b) show non-subject coding properties, they undergo deletion by Equi, undergo and control reduction under identity with a nominative subject, and undergo indefinite subject deletion, all properties generally restricted to subjects:[13]

(36) Accusative experiencers delete under Equi
Him burþ to liken well his lif
him befits to V[acc] will his life(nom)
'It befits him to like his life well.' (Dame Sirith: 82, A.D. 1275)

(37) Accusative experiencers undergo conjunction reduction under identity with nominative subject

a. Lewed men leued hym well and liked his wordes
'Ignorant men loved him well and his words pleased [them].'
(Langland, Piers Plowman, prol. 72, A.D. 1362)

b. Arthur loked on the swerd and liked it passynge well
(Mallory, Morte D'Arthur, A.D. 1470)

(38) Accusative experiencers control deletion by conjunction reduction of a nominative subject.

Us sholde neither lakken gold ne gere, But ben
(acc) should neither V[acc] gold nor gear but be

honoured while we dwelten there
honoured while we dwelt there

'We should not lack gold nor gear, but be honoured while we dwelt there.' (Chaucer, Canterbury Tales)

(39) Accusative experiencers undergo indefinite subject deletion

Good is quaþ Iosef, to dremen of win
'[It] is good, said Joseph, to dream of wine.' (Genesis and Exodus:2067,A.D.1250)

The above examples, and possibly others cited by Gaaf and Harris, show that experiencers in English exhibited behavioral subject properties before they exhibited coding properties.

The West Germanic data, taken in isolation, might be analyzed in either of two ways, which we shall label Hypothesis One and Hypothesis Two. According to Hypothesis One, experiencers and passivized dative objects displayed no subject properties in the parent language. In English, behavioral properties were acquired by some experiencer NPs by the thirteenth century, and coding properties later. But in German,

neither behavioral nor coding properties were acquired.[14] Thus, German is considered to have preserved the situation found in the proto-language. This is the position for which we shall argue.

According to Hypothesis Two, in the proto-language, as a result of the loss of subject coding properties at an even earlier stage, behavioral but not coding properties were associated with experiencers and passivized dative objects. In German these behavioral properties were lost, but in English they were not. At a later stage in English, these noun phrases re-acquired subject coding properties.

Although taken in isolation, the two hypotheses just outlined might appear to be equally plausible, when viewed in conjunction with the North Germanic and East Germanic data discussed earlier, only Hypothesis One is tenable. It will be remembered that in Gothic, the nominative NP in experiencer construction, rather than the dative experiencer, exhibited subject behavioral properties. Moreover, in Old Icelandic dative experiencers had only incipient subject behavioral properties, and passivized non-accusative objects appear to have had none. A clear trend toward the development of additional behavioral and, later, coding properties was seen in the case of Icelandic. This trend is consistent with the claim that the noun phrases in question showed no subject properties in the proto-language, but is is inconsistent with the hypothesis that in the proto-language a full panoply of subject behavioral properties were present. Hence, we conclude that the most plausible reconstruction is that in the proto-language experiencers and passivized non-accusative objects had no subject properties. Subsequently, first behavioral and then coding properties were developed in some of the daughter languages. Thus, the most plausible reconstruction for Germanic supports our hypothesis regarding the order of acquisition of behavioral and coding subject properties.

III. Polynesian

Further evidence in support of our hypothesis is provided by the existence of morphologically ergative languages in which transitive subjects are assigned historically oblique case markers. In a number of such languages, subjects marked in this way appear to have been historically oblique NPs (Stage A), which acquired the behavioral properties of subjects while retaining their oblique case marking (Stage B), with subsequent reinterpretation of the oblique marking as the normal marking for transitive subjects. In this section, we will trace a development of this type in Polynesian, as discussed in Chung (1976). Additional possible instances of this sort of change are treated by Anderson (1977).

The Polynesian languages fall into two large groups, those with nominative-accusative case marking and those with ergative-absolutive case marking. In the nominative-accusative languages, the subjects of active transitive verbs are unmarked. Upon passivization, those subjects are assigned the marker e. Compare the following examples from Maori:

(40) Active sentence in Maori

ka patu i te tuna te kōtiro (Chung, 17c, p.16)
aor. hit acc the eel the girl
'The girl killed the eel.'

(41) Passive sentence in Maori

ka tuku-a e Paowa te waka (Chung,19a,p.18)
aor. leave-pass agt Paowa the canoe
'The canoe was left by Paowa.'

In the ergative group, the subjects of active transitive verbs are marked with e. Observe the following example from Tongan, a language of the Tongic subgroup of Polynesian:

(42) Transitive sentence in Tongan

na'e taa'i 'e Mele 'a Sione (Chung, 22c,p.20)
past hit agt Mary abs John
'Mary hit John'

Thus, e is used for both passive agents in Maori and transitive subjects in Tongan. Chung argues convincingly that Proto-Polynesian was a nominative-accusative language, like Maori, and that the ergative case marker e in languages like Tongan developed from an original passive agent marker of the sort represented in the Maori example in (41). Chung (1976: 316) represents this development in the following way:

(43) Passive: Verb-Cia[15] e Agent (=underlying subject) Subject (=underlying DO)

reinterpreted as

Transitive: Verb-Cia e Subj DO

We will not repeat her extensive arguments for such a development in the present paper. Instead, we will accept the correctness of those arguments, and will demonstrate the way in which the development represented in (43) supports our hypothesis that behavioral properties are acquired prior to coding properties.

In the nominative-accusative languages representative of Stage A, the e NP (the underlying subject) in the passive sentences does not exhibit any of the behavioral properties of subjects. Maori, for example, has a rule of Equi which deletes those subjects which are also agents, as in (44):

(44) Equi deletes agentive subjects

i hiahia au ki te patu i te poaka (Chung,44a,p.127)
past want I comp hit acc the pig
'I wanted to kill the pig.'

The e NPs in passive sentences cannot be deleted even though they are agents:[16]

(45) Equi does not delete passive agents

*i hiahia au ki te patu-a te poaka (Chung,57a,p.130)
past want I comp hit-pass the pig
('I wanted to kill the pig.')

Maori also has a rule of subject-to-subject raising, governed by sentential negatives like kāhore 'not', which can apply to subjects, but not to direct objects or other non-subject NPs. Compare (46a) and (46b):

(46) a. Subjects undergo SSR

kāhore a Hōne i patu i te poaka (Chung,158a,p.179)
not pers John past hit acc the pig
'John didn't kill the pig.'

b. Non-subjects do not undergo SSR

*kāhore te poaka i patu ai a Hōne (Chung,155a,p.178)
not the pig past hit pro pers John
('John didn't kill the pig.')

The rule can apply to the unmarked NP (the underlying DO) in passive sentences, but not to the e NP (the underlying subject). Compare (47a) and (47b):

(47) a. Passive subjects undergo SSR

kāhore rāua kia kite-a mai e ngā tāngate
not they sbj see-pass here agt the man
o te whare (Chung, 159a, p.179)
of the house
'They were not seen by the people in the house.'

b. Passive agents do not undergo SSR

*kāhore (e) ngā tāngata kis kite-a (ai) rāua
not agt the-pl man sbj see-pass pro they-dual
('They weren't seen by the people.') (Chung, 156a, p.178)

Thus, at Stage A, as represented by Maori, underlying subjects in passive sentences in Polynesian lacked both the coding properties and the behavioral properties of subjects; passivized underlying direct objects, on the other hand, exhibited both subject behavioral properties and the zero case marking of subjects.

In the innovating, morphologically ergative languages of Polynesian, the e-marked NP (the historical agent phrase) in transitive sentences is treated as subject by such rules as Equi and Raising, while

the historical passive subject is not. Tongan, for example, has a rule of Equi which can delete intransitive or middle subjects, but not non-subject NPs. Compare the examples in (48):

(48) a. Equi deletes subjects

pea na'e 'alu 'a e tangatá 'o folau mama'o
and past go abs the man comp sail far
'Then the man went and sailed away.' (Chung, 18a, p.114)

b. Equi does not delete non-subjects

*na'e 'alu 'a e tamaiki tangatá 'o ngalo
past go abs the children man comp forgotten
'enau lesoni (Chung, 25a, p.117)
their lessons
('The boys went and forgot their lessons.') (deletion of stative agent)

In transitive sentences, the e NP can be deleted by Equi, but the historical passive subject cannot. Compare the examples in (49):

(49) a. Historical passive agents delete under Equi

na'e 'alu 'a e tangatá 'o taa'i 'a e kulī
past go abs the man comp hit abs the dog
'The man went and hit the dog.' (Chung, 26a., p.117)

b. Historical passive subjects do not delete under Equi

*na'e 'alu 'a e fefiné 'o u'u 'e he kulī
past go abs the woman comp bite erg the dog
('The woman went and dog bite Ø.') (Chung, 23a, p.116)

Thus, Equi in Tongan treats the historical oblique agent, not the historical passive subject, as subject.

Tongan also has a rule of subject-to subject raising, governed by the verb lava 'be able', which can raise intransitive subjects, but not non-subject NPs. Compare (50a) with (50b):

(50) a. Subjects undergo SSR
'e lava 'a e pēpē 'o lea
unm can abs the baby comp talk
'The baby can talk' (Chung, 63a, p.136)

b. Non-subjects do not undergo SSR

*te mau lava pē 'o ngalo ia
unm we can emp comp forgotten it
('We might be able to forget about it.') (Chung, 86b, p.145)
(raising of stative agent)

In transitive sentences in Tongan, subject-to-subject raising treats the e-marked NP, but not the historical passive subject, as subject. Compare (51a) and (51b):

(51) a. Historical passive agents undergo SSR

'e lava 'e ho fāmilí 'o tauhi koe
unm can erg your family comp care you
'Can your family take care of you?' (Chung, 87c, p.146)

b. Historical passive subjects do not undergo SSR

*na'e lava 'a e fefiné 'o taa'i 'e Sione
past can abs the woman comp hit erg John
('The woman could be hit by John.') (Chung, 84a, p.145)

In Tongan, therefore, historically oblique passive agents have taken on the behavioral properties of subjects while retaining oblique case marking. Correspondingly, original passive subjects have lost behavioral subject properties, but are still marked in the same way as intransitive subjects. These facts are consistent with the prediction of our hypothesis, which holds that subject coding properties will not be acquired by an NP prior to subject behavioral properties.

A similar situation exists in Samoan, another morphologically ergative language, in which historically oblique passive agents with e have come to behave as surface subjects, while historically passive subjects have ceased to behave as subjects. Samoan has a rule of subject-to-subject raising, governed by a number of verbs, including mafai 'be able', which can raise intransitive subjects, but not non-subject NPs. Compare (52a) and (52b):

(52) a. Subjects undergo SSR

e mafai e tātou 'ona nonofo
unm can erg we comp stay
'We can stay.' (Chung, 97a, p.151)

b. Non-subjects do not undergo SSR

*e mafai (e) le lā'au 'ona 'ou tā-ina 'oe
unm can erg the stick comp I hit-trans you
('I can hit you with the stick.') (Chung, 118b, p.160)
(raising of instrument)

As in Tongan, Samoan subject-to-subject raising can operate on historical oblique passive agents, but not on historical passive subjects Compare the examples in (53):

(53) a. Historical passive subjects do not undergo SSR

*e mafai le tali 'ona maua e le tama
unm can the answer comp catch erg the boy
('The boy can come up with the answer,') (Chung, 116a, p.159)

(53) b. Historical passive agents undergo SSR

e mafai e le tama 'ona maua le tale
unm can erg the boy comp catch the answer
'The boy can come up with the answer.' (Chung, 119a, p.160)

In summary, at Stage A in Polynesian, which is still represented by Maori, there was a productive rule of passive, which made underlying direct objects into subjects, and demoted underlying subjects to the role of oblique NP. At this stage, the e-marked underlying subject in passive sentences exhibited neither the behavioral nor the coding properties of subjects, while the underlying direct object exhibited both coding and behavioral properties of subjects. At Stage B, represented by Tongan and Samoan, the rule of passive has become non-productive, and the behavioral properties of subject-hood have been transferred from the historical passive subject to the historical agent phrase. Both NPs, however, preserved their original case marking, so that the historical passive subject, which is no longer a subject behaviorally, bears the same case as the subjects of intransitive sentences, while the historical oblique agent, which is a subject behaviorally, bears oblique case marking.

Stage C, the acquisition of subject coding properties by historically oblique passive agents, does not appear to have been reached in any of the Polynesian languages. In other language families, however, in which morphological ergativity arose through a development similar to that found in Polynesian, there is evidence for Stage C. Comrie (1977:p.38) cites the case of Indo-Iranian, in which passive constructions became dominant in the perfect, to the exclusion of active constructions, with the consequent loss of a productive passive in the perfect. Subsequently, behavioral subject properties were transferred from the historical passive subject to the historical oblique agent phrase. The former, however, retained control of verb agreement, and the zero case marking characteristic of subjects in other tenses, as illustrated by the following example from Hindi:

(54) Absolutive object, not ergative subject, controls agreement

is larke ne pustak parhī
this boy erg book(f abs) read(f)
'This boy read a book.' (Comrie, 139, p.39)

Comrie observes, however, that in some Indo-Iranian languages, control of agreement is being assumed by the historical passive agent, and that in Modern Persian, the historical passive agent has acquired all subject properties. Thus, the development and subsequent loss of ergativity may illustrate the three stages of the acquisition of subjecthood.

IV. Georgian

So far, we have examined the acquisition of subject properties in Polynesian and Germanic. In both cases there is good reason to believe that behavioral properties were acquired prior to coding properties. Although stage A is not directly attested in either case, there is

strong motivation for reconstructing a stage in which the relevant noun phrases lack any subject properties.

We shall now turn to Georgian. Our discussion is based on Harris (1976), Tschenkeli (1958) and Braithwaite (1973). We shall consider the status of dative experiencers like gelas in (55):[17]

(55) Dative experiencers

gelas	uqvars	nino
Gela(dat)	him-loves-she	Nino(nom)

'Gela loves Nino.'[18]

In both Old Georgian and Modern Georgian, the dative noun phrase in (55), gelas, exhibits a number of behavioral subject properties. A change has taken place in Modern Georgian which shows that these noun phrases are taking on morphosyntactic subject properties. Hence, data from Georgian corroborate our claim that coding subject properties are acquired subsequent to behavioral subject properties.

We shall turn first to the behavioral subject properties of dative experiencers. Examples are given from Modern Georgian and Old Georgian. There are two relevant tests of behavioral subject properties. The first test for subjecthood which we shall consider is tav- reflexivization.

As was shown by Harris, tav- may be coreferential only to the subject of its own clause. This is illustrated for sentences with nominative subjects in (56)-(57):

(56) Direct objects do not control reflexivization

mxaṭvari	daxaṭavs	vanos	tavistvis
painter(nom)	he-paints-him	Vano(dat)	self-for

'The painter$_i$ will paint Vano$_j$ for himself$_i$.' (Harris,1:2)

(57) Indirect objects do not control reflexivization

nino	ǎcvenebs	paṭara	givis	tavis
Nino(nom)	she-shows-him-him	little	Givi(dat)	self's
tavs	sarkeši	(Harris 1:3a)		
self(dat)	mirror-in			

'Nino$_i$ is showing little Givi$_j$ herself$_i$ in the mirror.'

Sentence (56) shows that tav- can refer only to subjects and not to direct objects. Similarly, in (57) tav- may only refer to the subject Nino, and not to the indirect object Givi.

Turning to verbs taking dative experiencers, we see that it is the dative which is the antecedent of tav-. This is shown in Modern Georgian in (58)-(60), and in Old Georgian in (61):

(58) Dative experiencers control reflexivization

temurs uqvars tavisi tavi
Temur(dat) him-loves-he self's self(nom)
'Temur loves himself.' (Harris 8:28a)

(59) Nominatives do not control reflexivization

*tavis tavs uqvars temuri
self's self(dat) him-loves-he Temur(nom)
('Himself loves Temur.') (Harris 8:28b)

(60) Initial nominatives do not control reflexivization

*temuri uqvars tavis tavs
Temur(nom) him-loves-he self's self(dat) (Harris,8:28c)

(61) Dative experiencers control reflexivization

cudad mogimedgrebies tavi
poorly you-give-up-it self(nom)
'You have given up on yourself badly.'[19] (A. Harris, personal communication)

Sentence (58) shows that the nominative may not be the antecedent. It should be noted that these facts cannot be due to word order. In (60) the nominative noun phrase precedes the dative, but the former still cannot serve as the antecedent for tav- reflexivization. Thus, dative experiencers are treated as subjects by tav- reflexivization.

We shall now consider a second rule that treats dative experiencers as subjects (rather than as indirect objects). This rule is causative clause union. Causative clause union is a rule, found in many languages, which maps biclausal structures like (62):

(62) NP_1 $CAUSE_S[NP_2$ $VP]_S$

onto simplex strucstures like (63):

(63) $_S[NP_1$ cause-verb $NP_2 \ldots]$

In Georgian, as in many other languages, though not all, when the complement verb is intransitive, the subject of that verb is a direct object in the output of clause union. This is illustrated in (64)-(65):

(64) Active intransitive

bavšvi iʒinebs
child(nom) he-sleeps
'The child is sleeping.' (Harris 5:14a)

(65) Complement subject of intransitive becomes direct object in causative clause union

dedam bavšvi daaǯina
mother(erg) child(nom) she-caused-sleep-him
'The mother got the child to sleep.' (Harris 5:16a)

(Direct objects receive nominative case with certain verbs and tenses. See Harris, Chapter 1, for details.)

In contrast, the subjects of transitive verbs are indirect objects in the output of the rule. This is shown in (66)-(67):

(66) Active transitive

mziam cecxli daanta
Mzia(erg) fire(nom) she-lit-it
'Mzia lit the fire.' (Harris 5:17a)

(67) Complement subject of transitive becomes indirect object in causative clause union

mamam mzias daantebina cecxli
father(erg) Mzia(dat) he-caused-light-her-it fire(nom)
'Father made Mzia light the fire.' (Harris 5:19a)

Indirect objects of intransitive verbs remain indirect objects, as is seen in (68)-(69). (The effect of the rule on indirect objects of transitives is not relevant to the present discussion.)

(68) Active intransitive with indirect object

ḳomisia vanos eṣṭumra
commission(nom) Vano(dat) it-visited-him
'The commission visited Vano.' (Harris 5:20b)

(69) IO remains IO in causative clause union

direktorma ḳomisia vanos aṣṭumra
director(erg) commission(nom) Vano(dat) he-caused-visit-him-it
'The director had the commission visit Vano.' (Harris 5:22b)

We have seen that grammatical relations in the output of clause union are determined by the transitivity of the complement clause and by the grammatical relations of the noun phrases in the complement in the input to the rule. Thus, with intransitive sentences with dative experiencers like (70),

(70) Dative experiencer

ṭusaǯs šioda
prisoner(dat) him-hungered
'The prisoner was hungry.' (Harris 8:38a)

if the dative nominal is a subject in the input to clause union, it is predicted that it will be a direct object in the output of the rule. But if the dative nominal is an indirect object, it should remain an indirect object subsequent to clause union. (Compare (68)-(69).) As is seen in (71), dative nominals in sentences like (70) become direct objects:

(71) Dative experiencer becomes direct object in causative clause union

ṭusaɣi moaǰives
prisoner(nom) pro-cause-hunger-him
'They starved the prisoner.'/'They let the prisoner go hungry.'

Therefore, we conclude that dative experiencers are treated as subjects and not indirect objects by causative clause union.

We have shown that dative experiencers behave in a manner typical of subjects rather than of indirect objects. This is true not only in Modern Georgian but in Old Georgian as well. We shall now turn to coding properties. In Old Georgian these nominals displayed no subject coding properties. Rather, they manifest coding properties typical of indirect objects. In Modern Georgian, however, a change appears to have taken place. Dative experiencers behave as subjects rather than as indirect objects with respect to certain coding properties.

There are three relevant coding properties: case marking, person agreement, and number agreement. With respect to case marking and person agreement, in sentences like (55) (repeated),

(55) Dative experiencers

gelas uqvars nino
Gela(dat) him-loves-she Nino(nom)
'Gela loves Nino.'

the nominative noun phrase, nino, manifests subject coding properties in both Old Georgian and in Modern Georgian. Note that nino is in nominative case and triggers subject person agreement on the verb. The situation with regard to number agreement, however, is more complex. Verbs agree in number with any first or second person subject, direct object, indirect object. But they agree in number with third person noun phrases only when they are subjects. Third person number agreement is illustrated in (72)-(73):

(72) Number agreement (Third person subject)

a. darča
he-remained
'He remained.' (Harris 9:4a)

b. darčes
they-remained
'They remained.' (Harris 9:5a)

(73) Number agreement (Third person non-subject)

a. icnobs čems studentebs
he-knows-it my students(dat)
'He knows my students.' (Harris 9:6a)

b. ačukebs tavis c̣igns čems studentebs
he-gives-him-it self's book(dat) my students(dat)
'He is giving his book to my students.' (Harris 9:6b)

Thus, number agreement distinguishes between third person subjects and third person non-subjects. Only the former govern number agreement.

We shall now consider the status of the two noun phrases in sentences like (55). Does the verb agree with the nominative noun phrase or with the dative experiencer? In Old Georgian (Tschenkeli: 1958, 453-4) the verb always agreed in number with the nominative noun phrase.[20] This is shown in (74). Note the plural agreement on the verb.

(74) Nominative governs number agreement exclusively

me miqvaran isini
me(dat) me-love-they they(nom)
'I love them.' (Tschenkeli: p.454)

In Standard Modern Georgian a change has taken place. When both the dative experiencer and the nominative noun phrase are third person, the verb agrees with the dative rather than the nominative. Contrast (75), in which the first person nominative controls number agreement and not the third person dative, and (76), in which the third person dative, and not the third person nominative, controls verb agreement:

(75) First person nominative controls number agreement

a. mas vuqvarvar me
him(dat) him-love-I I(nom)
'He loves me.'

b. mas vuqvarvart cven
him(dat) him-love-we we(nom)
'He loves us.'

c. mat vuqvarvar me
them(dat) him-loves-I I(nom)
'They love me.'

d. mat vuqvarvart čven
them(dat) him-loves-we we(nom)
'They love us.' (Tschenkeli: p.459)

(76) Third person dative and not third person nominative controls number agreement

a. mas uqvars is
him(dat) him-loves-he he(nom)
'He loves him.'

b. mas uqvars isini
him(dat) him-loves-they they(nom)
'He loves them.'

c. mat uqvart is
them(dat) them-love-he he(nom)
'They love him.'

d. mat uqvart isini
them(dat) them-love-they they(nom)
'They love them.' (Tschenkeli : p.459)

Hence, in Standard Modern Georgian, the dative experiencer has acquired partial control of number agreement, a coding property.

In colloquial Modern Georgian, the acquisition of control of number agreement by the dative experiencer has, according to Tschenkeli (1958), been generalized. While in Standard Modern Georgian control of agreement by third person datives is limited to environments in which the nominative noun phrase is also third person, in colloquial Modern Georgian, this has been generalized for some speakers to environments in which the nominative noun phrase is first or second person. Thus, in (77) (from Harris (1976)), the dative experiencer triggers plural agreement despite the presence of a first person pronoun in the clause.

(77) Dative governs number agreement regardless of person of nominative

mšobleb-s vuqvarvart (me)
parents(dat) them-love-I I(nom)
'My parents love me.' (Harris p.283 fn.5)

This example illustrates the extension of the control of number agreement by dative experiencers in environments not possible in Standard Modern Georgian.

We have shown that in Old Georgian behavioral subject properties are controlled by the dative experiencer and coding properties by the nominative noun phrase. In Modern Georgian, especially in the colloquial language, one coding property, number agreement, has come to be controlled by the dative experiencer. Hence, Georgian corroborates our claim that nominals acquiring subject properties first acquire behaviroal properties and only later coding properties.[21]

V. Conclusions

We have examined the historical process by means of which, in a number of languages, subject properties were acquired by a noun phrase not previously having those properties. Among the languages examined were Polynesian (Maori, Tongan, and Samoan), Germanic (Gothic, Icelandic, Faroese, Swedish, German, and English), and Georgian.

From the languages surveyed, a common pattern emerges. There are three stages in the acquisition of subject properties. In the first stage, the noun phrase in question shows no subject properties. In the second stage, the noun phrase begins to take on transformational properties associated in the language with subjecthood. During the third stage, subject coding properties are acquired. Among the coding properties are control of such morphosyntactic processes as case marking and verb agreement. What distinguishes the third stage from the second is that the syntactic properties acquired during the third stage are reflected in the morphology of the language. Today's morphology, as Givón (1971: 413) and others have noted, is the ossification of yesterday's syntax. Hence, it is not surprising that morphosyntactic processes are acquired later than purely transformational processes. It is just such an order that we have observed.

NOTES and REFERENCES

*The idea for this paper was suggested by the work of Yael Ziv (1976), who insightfully recognized the diachronic implications of Keenan's (1976) claims regarding the synchronic acquisition of subject properties in a derivation. A number of linguists offered us valuable comments and supporting data. These include Alice Harris, who supplied crucial data on Georgian (as well as helpful discussion of the data), and Anthony Britti, Sandy Chung, Bernard Comrie, Dee Holisky, Hans H. Hock, Janice L. Jake, Jerry Morgan, Alan Timberlake, Rochelle Wright, and Ladislav Zgusta, who provided comments on the data and/or the analysis.

We should also like to thank the National Science Foundation (grants SOC 75-00244 and BNS 77-27159) and the Research Board of the University of Illinois for supporting this research. Earlier versions of this paper were presented as a Forum Lecture at the 1977 Linguistic Institute in Hawaii, as well as at the 1977 Summer Meeting of the LSA, the Third International Congress on Historical Linguistics (Hamburg), the linguistics colloquia of Cambridge University and Tel Aviv University, and at the Seventh All-India Conference of Dravidian Linguists, Bangalore, June 1977.

[1]There are two possible ways of conceptualizing the changes discussed in this paper. We are claiming that certain originally non-subject NPs are becoming subjects historically, and that they are therefore coming to be treated as such by rules which are sensitive to subjecthood. Alternatively, it might be claimed that the rules themselves

are changing, starting to apply not only to subjects, but to certain non-subject NPs (e.g. experiencers) as well. We reject the latter view primarily because of its unnecessary complexity. It is generally the case that if one rule of a given type (e.g. a behavioral rule) treats such NPs as it treats subjects, others of the same type will treat them similarly. The second of the above positions would require the positing of several changes in individual rules in order to account for this fact. The position that it is the NPs which are changing in grammatical role accounts for these parallel extensions of the applicability of individual rules as the consequence of a single change. Moreover, the rule-change position unnecessarily requires the weakening of the otherwise valid claim that certain processes, e.g. deletion by Equi, are cross-linguistically restricted to subjects (cf. Keenan (1976) inter alia).

[2]For a fuller list of behavioral and coding properties, see Keenan (1976). The distinction between behavioral and coding properties is based on Keenan's paper. Our paper provides additional support for this distinction.

[3]Keenan (1975: p.324) lists sentence-initial position as a specific coding property of subjects. Even if this claim is assumed to be correct, and if initial position can be shown to be acquired prior to behavioral subject properties, our major contention, that there is a language-independent order of acquisition of the form: (1) behavioral subject properties, and (2) all coding properties aside from word order, remain unaffected.

[4]Although we refer to these nominals as "passivized non-accusative objects" for all the Germanic languages, we should like to emphasize that for languages like German in which these NPs have acquired no subject properties, this terminology is somewhat misleading. In the case of Stage A, the term "passivized non-accusative objects" should only be taken as a means of identifying those nominals that at a later stage come to demonstrate derived subject properties in the output of passivization.

[5]The question of the internal grouping of the Germanic languages has not been settled with certainty. Various investigators have advocated a tripartite division into North, West, and East Germanic, a bipartite division into West and Northeast Germanic, and a division into East and Northwest Germanic. The arguments for these divergent positions have been reviewed in Kufner (1972). This issue does not affect our argument.

[6]The evidence regarding the acquisition of behavioral and coding properties by passivized non-accusative objects in Gothic is conflicting. In most instances, such objects seem to have acquired properties of both types. That is, they seem to have become true subjects. Compare the following examples:

(i) ei [eis] in filuwaurdein seinai andhausjaindau
that they(nom) in verbosity refl. V[dat] pass 3pl
'that they will be heard in their verbosity' (Mat 6:7)

(ii) gaweisodai waurþun daurawaurdos
V[gen] part aux 3pl porters (nom)
'Porters were appointed.' (Neh 7:1)

Note that in (i), the passivized dative object not only appears as a nominative and controls verb agreement, but serves as the controller of reflexivization as well. In all such cases, however, the Gothic construction translates a normal Greek passive construction with a nominative subject. This suggests the possibility that the Gothic construction has been influenced by the Greek, especially in view of the fact that in the few instances in which the Gothic NPs did not acquire coding properties, the Gothic construction was either independent of, or in opposition to the Greek, as in (iii) and (iv):

(iii) þairh þatei is brukjaidau
through that it V[gen] pass
'through that it is used'/'through its use' (Col 2:22)

(iv) jah bajoþum gabairgada
and both (dpl) V[dat] pass sg
'And both will be saved.' (Mat 9:17)

[7]There are instances in which the reflexive possessive adjective refers to dative experiencers with verbs other than þykkia. Rose (1976: 128) cites one example in particular involving the verb sýna 'seem'. She observes, however, that the reflexive possessive adjective, unlike reflexive pronouns, is used to refer to a variety of non-subject NPs, as in the following example:

(i) hann skyldi fá henni eign sína
he should get her(dat) possessions refl.
'He should get her her possessions.' (Heimskringla I:422)

Thus, the ability of an NP to serve as the antecedent for a reflexive possessive adjective in Old Icelandic does not constitute a reliable subject property.

[8]There are a number of attestations in Old Icelandic in which experiencer predicates appear embedded under subject-to-object raising verbs, as in the following examples:

(i) Grettir kvaþ sér þat betr þykkia
Grettir said her(dat) that better seem (inf)
'Grettir said that to seem better to her.' (Grettis Saga, 39)

(ii) Grettir kvaþ sér þat allvel líka
Grettir said her(dat) that well like (inf)
'Grettir said that to please her well.' (Grettis Saga, 78)

(iii) Hrólfr kua-zk ekke suá þykkia
Hrolf(nom) said-refl not so seem (inf)
'Hrolf said [it] not to seem so to him.' (cited in Heusler, 1967: 39)

In (iii), the underlying experiencer pronoun has cliticized to the main clause verb. Unfortunately, it does not appear possible to determine whether in cases like the above the complement experiencer or the complement nominative has been raised by SOR into matrix object position. It will be seen below that in Modern Icelandic, when a dative experiencer undergoes SOR, it retains its dative case.

[9]See Andrews (1976: 13ff) for details.

[10]Passivized genitive objects, however, may not:

(i) *ég vonast til að verða beðið
I hope comp aux inf V[gen] supine
('I hope to be waited for.')

[11]The information provided by Lockwood on experiencers in Faroese is insufficient for determining to what extent these have undergone the B to C shift.

[12]The experiencers in these languages display behavioral as well as coding properties. In the following Swedish examples, nominative experiencer subjects which developed from historically non-nominative experiencers control reflexivization, delete through Equi, and are raised by subject-to-subject raising:

(i) han tyckte sig vara i en trevlig stad
he(nom) thought refl be in a pleasurable city
'He thought himself to be in a pleasurable city.'

(ii) jag hoppas drömma
I hope dream (inf)
'I hope to dream.'

(iii) han tycks minnas
he seems remember (inf)
'He seems to remember.'

[13]The examples in (37) are cited by Harris as examples (1b) and (1d) in the appendix to her paper. Our example (38) was cited by Butler (1976). We have used examples (34) and (36) rather than similar ones cited by Gaaf (1904: 32) and Harris (1975: 20) for chronological reasons. The examples which they cite are of uncertain significance, since they postdate the first occurrence of the verbs involved with nominative experiencers, and since nominative experiencers appear with these verbs elsewhere in the works of the authors cited. Our examples (34) and (36) present no such problem, since they substantially antedate the first occurrences of these verbs with subject-coded experiencers.

[14]Some speakers of Austrian German allow passivized non-accusative objects to be deleted by Equi. For these speakers, sentences like (23) are grammatical. Thus, such NPs seem to be acquiring some behavioral properties in some dialects of German.

[15]C represents a set of consonants whose occurrence in the passive suffix is lexically determined.

[16]Also undeletable are the underlying DOs of passive sentences, since these do not meet the agency condition of the application of Equi, as in the following example:

(i) *i hishia au ki te patu-a e Rewi
past want I comp hit-pass agt Rewi
('I wanted to be hit by Rewi.) (Chung, 54a, p.150)

[17]Georgian also has a special mode described by Harris as "evidential" (p. 226). This is the mode used where a speaker reports an event which he did not directly witness, but of which he saw the results of, or some evidence that it occurred. Using this mode also relieves the speaker of direct responsibility for the accuracy of the statement and Harris uses turme 'apparently' to gloss this mode. This mode usually appears in a perfective tense. Thus, certain verbs show up in a nominative-subject/dative-DO and IO pattern in the present, but take the dative case for the subject and nominative for DO (and a postpositional phrase for the IO) in the evidential mode. Thus, in (i) and (ii) the same relations hold between the berb and the NPs. The only difference is in mode and in case marking:

(i) glexi tesavs teslebs
peasant(nom) he-sows-it seeds(dat)
'The peasant is sowing seeds.'

(ii) turme glexs dautesavs teslebi
apparently peasant(dat) him-sowed-it(nom) seeds(nom)
'The peasant has apparently sown seeds.'

Though we will primarily discuss the status of the experiencer dative in this paper, all the facts hold also for evidential mode datives, and we shall occasionally use examples from this construction.

[18]Harris glosses this sentence as:

(i) gelas uqvars nino
Gela(dat) he-loves-her Nino(nom)
(he=Gela, her=Nino)

However, in our gloss we have, for the sake of clarity, decided to use nominative pronouns for those markers which agree with nominative nouns, and dative pronouns for those agreeing with dative-nominals. This is closer in spirit to what actually happens in Georgian: the markers referring to the dative are object markers and those referring to the nominative are subject markers. We do not give the markers and the verb root

in the same order in which they appear in Georgian (where the object marker usually precedes the root and the subject marker follows the root) due to the fact that the markers are sometimes discontinuous (i.e., circumfixes) and sometimes ∅ morphemes.

[19]The verb can also mean 'weaken'. In the verb form, -g- is the second person (singular) object marker: -es is the third person singular subject marker, agreeing with tavi. The verb is in the perfect or first evidential, which is part of Series III. The initial subject is 'you', which undergoes inversion and becomes the final indirect object, triggering indirect object agreement and occurring in the dative case. See fn. 17 for a brief discussion on inversion (in the evidential mode).

[20]The same facts are reported by Braithwaite (1973: 110) (our glossing):

Old Georgian (i) mas mohuklavs igi
him(dat) him-killed-he he(nom)
'Apparently, he has killed him.' (-s marks 3sg)

(ii) mas mohuklavan igini
him(dat) him-killed-they they(nom)
'Apparently, he has killed them.' (-an marks 3pl subject)

Here we see that in Old Georgian the derived subject (Nom) triggers third person number agreement. In Modern Georgian, the Nom does not have control over third person number agreement anymore and as a result, even though we have a plural Nom NP, the verb is marked for third person singular subject in agreement with the dative NP, as observed in (76b).

[21]Georgian differs from Germanic in that we have no evidence for a stage in which dative experiencers show no syntactic subject properties. We hypothesize, however, that there was such a stage.

REFERENCES

Anderson, Steven R. 1977. On the mechanisms by which languages become ergative. In C. Li, ed.: Mechanisms of syntactic change. Austin, Texas: Texas Press.

Andrews, Avery. 1976. The VP complement in Modern Icelandic. In Montreal working papers in linguistics, vol. 6.

Bennett, J.A.W. and G.U. Smithers (eds.). 1966. Early Middle English verse and prose. Oxford: Clarendon Press.

Braithwaite, Kim. 1973. Case shift and verb concord in Georgian. Dissertation, University of Texas at Austin.

Breckenridge, Janet. 1975. Rules which nothing undergoes. Bachelor's thesis, Harvard University: Cambridge, Mass.

Butler, M.L. 1976. The reanalysis of Middle English impersonal constructions and the characterization of 'subject of a sentence.' In Carlota S. Smith and Susan F. Schmerling, (eds.): Texas Lingustic Forum, vol. 4. Austin.

Chung, S.L. 1976. Case marking and grammatical relations in Polynesian. Dissertation, Harvard University: Cambridge, Mass.

Comrie, B. 1977. Ergativity. Mimeo.

Ebert, R. 1975. Subject raising, the clause squish, and German scheinen constructions. CLS 11.

Evans, E.E. 1975. Psychological process experiencer verb impersonals: a case grammar approach. Dissertation, University of Texas.

Falk, Hjalmar and Alf Torp. 1900. Dansk-Norskens syntax. Kristiania H. Aschehoug.

Gaaf, W. van der. 1904. The transition from the impersonal to the personal construction in Middle English. Heidelberg.

Givón, Talmy. 1971. Historical syntax and synchronic morphology: an archaelogist's field trip. CLS 7.

Harris, Alice. 1975. On the mechanism of syntactic rule loss: the evidence from English and Udi. Unpublished paper.

_____1976. Grammatical relations in Modern Georgian. Dissertation, Harvard University: Cambridge, Mass.

Heusler, Andreas. 1967. Altisländisches Elementarbuch. 7th edition. Heidelberg: Carl Winter.

Keenan, E. 1976. Towards a universal definition of 'subject'. In C. Li, (ed.): Subject and Topic. Academic Press.

Kufner, Herbert L. 1972. The grouping and separation of the Germanic languages. In Frans van Coetsom and Herbert L. Kufner, (eds.): Toward a grammar of Proto-Germanic. Tübingen: Max Niemeyer.

Lockwood, W.B. 1964. An introduction to Modern Faroese. Copenhagen: Munksgaard.

Nygaard, M. 1905. Norrøn syntax. Kristiania: H. Aschehoug.

Rose, M.L. 1976. The syntax of the reflexive pronoun in Gothic and Old Norse. Ph.D. Thesis, University of Wisconsin.

Streitberg, Wilhelm (ed.) 1971. Die gotische Bibel. 6th edition. Darmstadt: Wissenschaftliche Buchgesellschaft.

Tschenkeli, Kita. 1958. Einführung in die Georgische Sprache. Zurich: Amirani Verlag.

Ziv, Yael. 1976. On the reanalysis of grammatical terms in Hebrew possessive constructions. In P. Cole, (ed.): Studies in Modern Hebrew and Semantics, pp. 129-152. North-Holland Pub. Co.

Studies in the Linguistic Sciences
Volume 8, Number 1, Spring 1978

ACCUSATIVE SUBJECTS IN IMBABURA QUECHUA

Peter Cole and Janice L. Jake

This paper examines two impersonal constructions in Imbabura Quechua in which a noun phrase in accusative case displays a variety of syntactic subject properties. One impersonal constuction is governed by lexical, or non-derived, experiencer verbs. In the other construction experiencers are formed by adding the desiderative suffix -naya- to a verbal stem. Although in both constructions the accusative experiencer nominal possesses certain, though not all, properties of subjects, the two constructions differ with regard to which subject properties are displayed in each, indicating that the nominals in -naya- desideratives have reached a more advanced stage in reanalysis from non-subject to subject than have those with lexical verbs of experience.

1. Introduction

In Imbabura Quechua (IQ)[1], and in other varieties of Quechua as well, subjects typically appear in nominative case and control verb agreement. There are, however, two 'impersonal constructions in IQ in which a noun phrase in accusative case displays a variety of syntactic subject properties. The two constructions are illustrated in (1) and (2).

(1) Experiencer verbs[2]

a. can-da rupa-n
2sg-acc burn-3
'You are hot.'

b. ñuca-ta nana-n
1sg-acc hurt-3
'I hurt.'

(2) -naya- desideratives

a. can-da micu-naya-n
2sg-acc eat-desid.-3
'You would like to eat.'

b. ñuca-ta puñu-naya-n
1sg-acc sleep-desid.-3
'I would like to sleep.'

The sentences of (1) are instances of what we shall refer to as the lexical, or non-derived, experiencer construction. In addition to rupa- 'to be hot' and nana- 'to hurt' (intransitive), chiri- 'to be cold', yarja- 'to be hungry' and muna- 'to want' may be used in this

construction.[3] The sentences of (2) are -naya- desideratives. These[4] are formed by adding the desiderative suffix -naya- to a verbal stem.

The purpose of this paper is to sketch the morphological and syntactic properties of the nominals in each of the above constructions. We shall argue that in both constructions the 'experiencer' nominal (can-da and ñuca-ta in (1) and (2)) possesses certain, though not all, properties of subjects in IQ. Furthermore, the two constructions differ with regard to which subject properties are displayed in each. We shall show that the nominals in -naya- desideratives have reached a more advanced stage in reanalysis from non-subject to subject than have those with lexical verbs of experience.

Our discussion is organized as follows: In each section, we determine the extent to which the nominals in question are treated as subjects or as objects by grammatical processes found in IQ. Different groups of rules are considered in each section. In section 2 we examine the morphosyntactic properties of the experiencer nominals in the constructions with regard to the subject coding rules[5] of Case Marking (CM) and Verb Agreement (VA) and the object coding rule of Pronominal Object Cliticization (POC). Although experiencer nominals are typically coded as non-subjects by these rules, those in the desiderative construction may, and under certain conditions must, display partial subject coding properties. This is not the case for nominals with non-derived experiencer verbs like those of (1).

In section 3 we compare the nominals in the two constructions with regard to three rules: Equi NP Deletion (Equi), Subject Raising (SR) and Passivization. Nominals in both construction undergo the rules typically restricted to subjects, Equi, SR, and demotion by Passivization, and fail to undergo promotion to subjecthood by Passivization, this being restricted to objects.[6] Thus, the experiencer nominals in both constructions behave unambivalently as subjects with regard to the rules discussed in this section.

Two constructions in which subjects and non-subjects are treated differently are discussed in section 4. These are switch reference purpose and adverbial clauses. In these constructions, nominals with -naya- verbs behave unambivalently as subjects. Nominals with lexical experiencer verbs, however, may be treated either as subjects or as non-subjects.

In section 5, we consider the implications of the above facts for the diachronic and synchronic analysis of the two constructions. We argue that experiencer nominals with -naya- verbs are in the final stages of reanalysis as subjects. Nominals with lexical verbs of experience are at an earlier stage of reanalysis. This position is supported not only by data from IQ, but also by cross-linguistic evidence.

2. Morphosyntactic coding rules

In this section we shall consider the behavior of experiencer nominals in the two constructions with regard to three morphosyntactic coding rules: Case Marking, Verb Agreement and Pronominal Object

Cliticization. Experiencer nominals in both constructions behave as non-subjects with regard to VA. This is not true, however, with respect to CM or POC.

With the exception of the two constructions under consideration, subjects appear in nominative case (Ø case marking), control VA and fail to under POC. Direct objects receive accusative case (-ta), fail to control VA and may, under appropriate conditions, optionally undergo POC.[7] These facts are illustrated in (3)-(7):

(3) Subject in nominative case and controls VA

can-Ø ñuca-ta ricu-ngui
2sg-nom 1sg-acc see-2sg
'you see me.'

(4) Object cliticized

can-Ø (ñuca-ta) ricu-wa-ngui
2sg-nom 1sg-acc see-1-2sg
'you see me.'

(5) Object may not control VA

*can-Ø ñuca-ta ricu-ni
2sg-nom 1sg-acc see-1sg
(You see me.)

(6) Subject may not appear in accusative case

*can-ta ñuca-ta ricu-ngui
2sg-acc 1sg-acc see-2sg
(You see me.)

(7) Subject may not undergo cliticization

*ñuca-Ø can-da ricu-wa-ni
2sg-nom 2sg-acc see-1-1sg
(I see you.)

To return to experiencer constructions, with both desideratives and lexical verbs of experience the experiencer nominal is frequently coded as a superficial direct object. It receives accusative case, fails to trigger VA and, in the case of lexical experiencers, may undergo POC. This is shown in (8)-(11):

(8) -naya- experiencer in accusative and fails to constrol VA

ñuca-ta-ca can-da ricu-naya- {n / *ni}
1sg-acc-topic 2sg-acc see-desid- {3 / *1sg}
'I would like to see you.'

(9) Lexical experiencer in accusative and fails to control VA

(ñuca-ta) nana-{n / *ni}
lsg-acc hurt-{3 / *lsg}
'I hurt.'

(10) Lexical experiencer undergoes POC

a. (ñuca-ta) nana-wa-n
lsg-acc hurt-1-3
'I hurt.'

b. (ñuca-ta) chiri-wa-n
lsg-acc cold-1-3
'I am cold.'

Desiderative experiencers differ from lexical experiencers in that the desiderative experiencer nominal may optionally appear in nominative case and can never undergo POC. The option of nominative case marking does not exist for lexical experiencers.[8] Furthermore, desiderative experiencers differ from lexical experiencers in that the former fail to undergo the object coding rule of POC. This is illustrated in (11) (cf. 10):

(11) -naya- experiencer fails to undergo POC

a. *(ñuca-ta-ca) can-da ricu-naya-wa-n
lsg-acc-topic 2sg-acc see-desid-1-3
(I would like to see you.)

b. *micu-naya-wa-n
eat-desid-1-3
(I would like to eat.)

c. *puñu-naya-wa-n
sleep-desid-1-3
(I would like to sleep.)

Despite their appearance in nominative case, desiderative experiencers fail to control VA:

(12) -naya- experiencer in nominative case but fails to control VA

ñuca-Ø-ca can-da ricu-naya-{n / *ni}
lsg-nom-topic 2sg-acc see-desid-{3 / *lsg}
'I would like to see you.'

(13) Lexical experiencer may not appear in nominative

*ñuca-Ø-ca nana-{n / ni}
1sg-nom-topic hurt-{3 / 1sg}
(I hurt.)

We have shown that experiencers with -naya- verbs and experiencers with lexical verbs of experience are treated in different ways by morphosyntactic coding rules. -naya- experiencers may optionally be coded as subjects by CM and are obligatorily coded as subjects by POC (if POC applies). They fail, however, to trigger VA. In contrast, lexical experiencers are invariably treated as objects. As will be seen below, the two constructions differ in a similar way for certain, though not all, syntactic rules as well.

3. Some syntactic indications of subjecthood

In this section, we shall examine the behavior of experiencer nominals in both constructions with respect to three syntactic rules: Equi, SR, and Passivization. The ability of an NP to undergo these rules has been shown cross-linguistically to be indicative of the grammatical role of the nominal in question. Equi and SR take subjects as their input. The former rule deletes complement subjects under identity with a matrix noun phrase, while the latter promotes a complement subject to matrix direct object. In contrast, Passivization promotes underlying object noun phrases to subjecthood in a manner to be illustrated below.

For each rule, we shall first show that in IQ it is normally restricted to noun phrases with a certain grammatical role (subject for Equi and SR, object for promotion by Passivization, and subject for demotion by Passivization). We shall then show that experiencer nominals with both desiderative and lexical experiencer verbs behave like subjects rather than objects: They undergo deletion by Equi and promotion by SR; they undergo demotion by Passivization, but they fail to undergo promotion by Passivization. Thus, we claim, accusative experiencer nominals in both constructions exhibit significant subject behavioral properties.

Sentences (14)-(15) illustrate the application of Equi to sentences with nominative subjects:

(14) a. Input structure to Equi

pai-$Ø_i$ S[pai-$Ø_i$ trabaja-] gushta-n
3sg-nom 3sg-nom work like-3

b. Output of Equi

pai-$Ø_i$ $Ø_i$ trabaja-na-ta gushta-n
3sg-nom work-infin-acc like-3
'She likes to work.'

(15) a. Input structure to Equi

pai-$\emptyset_i$ $_S$[pai-$\emptyset_i$ mishqui-ta micu-] gushta-n
3sg-nom 3sg-nom candy-acc eat like-3

b. Output of Equi

pai-$\emptyset_i$ $\emptyset_i$ mishqui-ta micu-na-ta gushta-n
3sg-nom candy-acc eat-infin-acc like-3
'She likes to eat candy.'

Sentences (16)-(17) show Equi does not apply to objects:

(16) Complement object coreferntial to matrix subject

ñuca-$\emptyset_i$ muna-ni $_S$[pai-$\emptyset$ ñuca-ta$_i$ ricu-chun]
1sg-nom want-1sg 3sg-nom 1sg-acc see-subjunctive
'I want her to see me.'

(17) Equi fails to apply

*ñuca-$\emptyset_i$ muna-ni pai-$\emptyset$ $\emptyset_i$ ricu-na-ta
1sg-nom want-1sg 3sg-nom see-infin-acc
(I want her to see me.)

Equi can, however, delete the experiencer nominal in both the desiderative and the non-derived experiencer construction:

(18) a. Input structure to Equi with lexical experiencer

warmi-$\emptyset_i$ mana gushta-n-llu $_S$[warmi-ta$_i$ nana-ju-]
woman-nom not like-3-neg woman-acc hurt-prog

b. Output of Equi

warmi-$\emptyset_i$ mana gushta-n-llu $\emptyset_i$ nana-ju-na-ta
woman-nom not like-3-neg hurt-prog-infin-acc
'A woman doesn't like to hurt.'

(19) a. Input structure to Equi with -naya- desiderative

wira warmi-$\emptyset_i$ mana gushta-n-llu $_S$[wira warmi-ta$_i$ micu-naya]
fat woman-nom not gushta-n-llu fat woman-acc eat-desid

(19) b. Output of Equi

wira warmi-$\emptyset_i$ mana gushta-n-llu $\emptyset_i$ micu-naya-na-ta
fat woman-nom not like-3-neg eat-desid-infin-acc
'The fat woman doesn't like wanting to eat.'

Examples (18)-(19) show that unlike object nominals, experiencers in both of the constructions in question may be deleted by Equi. Thus, Equi treats the experiencer nominals in these constructions like subjects rather than objects.

The rule of SR also distinguishes between accusative experiencers on the one hand and objects of the other.[9] SR applies to nominative subjects and to -naya- and lexical experiencers. The examples of (20) show the effect of SR on nominative subjects:

(20) a. Input to SR

pai-$\emptyset$ S[ñuca-$\emptyset$ micu-j-ta] cri-rca[10]
3sg-nom 1sg-nom eat-pres. comp-acc believe-3past

b. Output of SR

pai-$\emptyset$ ñuca-ta cri-rca micu-j-ta
3sg-nom 1sg-acc believe-3past eat-pres. comp-acc
'He believed me to eat.'

c.

*pai-$\emptyset$ ñuca-$\emptyset$ cri-rca micu-j-ta
3sg-nom 1sg-nom believe-3past eat-pres.comp-acc
(He believed me to eat.'

d. Application of POC to (20b)

pai-$\emptyset$ cri-wa-rca micu-j-ta
3sg-nom believe-1-3past eat-pres. comp-acc
'He believed me to eat.'

In general, the application of SR can only be detected on the basis of word order (which is relatively free, and hence a poor test), and of the application of subsequent rules: e.g., case marking, matrix clause POC and matrix clause Passivization. Note that in (20b) the underlying complement subject is in accusative rather than nominative case. In (20d) this noun phrase has been cliticized to the matrix verb. These facts indicate that the underlying complement subject has become the derived matrix object.

Further indication that the underlying complement subject has become a derived matrix object is provided by Passivization The examples of (21) show that these nominals may be promoted to matrix subject by Passivization.

(21) a. Input to Passivization

Maria-Ø ñuca-ta cri-n $_S$[micu-shca-ta]
Maria-nom 1sg-acc believe-3 eat-past part-acc
'Maria believes me to have eaten.

b. Output of Passivization

ñuca-Ø-ca Maria-Ø cri-shca ca-ni
1sg-nom-topic Maria-nom believe-past part be-1sg

micu-shca-ta
eat-past part-acc

'I am believed by Maria to have eaten.'

The sentences of (22) show that SR and subsequent POC and Passivization are restricted to underlying complement subjects. Note that underlying complement objects may not undergo SR. This is demonstrated by the failure of these NPs to cliticize to the matrix verb or to be promoted to matrix subject by Passivization. (Accusative case marking is inapplicable as a test for SR because the complement object is marked accusative as a result of its grammatical role in the complement clause.)[11]

(22) a. Transitive object complement clause

José-Ø cri-n $_S$[Maria-Ø ñuca-ta ricu-shca-ta]
José-nom believe-3 Maria-nom 1sg-acc see-past part-acc
'José believes that Maria saw me.'

b. Complement object fails to cliticize to the matrix verb

*José-Ø cri-wa-n Maria-Ø ricu-shca-ta
José-nom believe-1-3 Maria-nom see-past part-acc
(José believes that Maria saw me.)

c. Complement object fails to become matrix passive subject

*ñuca-Ø-ca José-Ø cri-shca ca-ni
1sg-nom-topic José-nom believe-past part be-1sg

Maria-Ø ricu-shca-ta
Maria-nom see-past part-acc

(I am believed by José for Maria to have seen me.)

We have shown in the preceding paragraphs that SR distinguishes between complement subjects and complement objects. The former may, subsequent to SR, be cliticized to the matrix, or become the matrix passive subject. The latter may not. Thus, SR, in conjunction with POC and Passivization, provides a syntactic test for the status of -naya- and lexical experiencer nominals. If these nominals undergo

matrix cliticization and Passivization to matrix subject, we are justified in claiming that they are treated as subjects rather than objects by SR. We shall now show that this is, in fact, the case.

The sentences of (23) show that accusative -naya- experiencers may be cliticized to the matrix verb. This is shown for lexical experiencers in (24).

(23) a. Input to POC

jari-Ø ñuca-ta cri-n micu-naya-shca-ta
man-nom 1sg-acc believe-3 eat-desid-past part-acc
'The man believes me to have wanted to eat.'

b. Output of POC

jari-Ø cri-wa-n micu-naya-shca-ta
man-nom believe-1-3 eat-desid-past part-acc
'The man believed me to have wanted to eat.'

(24) a. Input to POC

jari-Ø ñuca-ta cri-n chiri-shca-ta
man-nom 1sg-acc believe-3 cold-past part-acc
'The man believes me to have been cold.'

b. Output of POC

jari-Ø cri-wa-n chiri-shca-ta
man-nom believe-1-3 cold-past part-acc
'The man believes me to have been cold.'

Similarly, experiencers in both constructions may undergo Passivization and become matrix subjects, as is seen for desideratives in (25) and lexical experiencers in (26).

(25) a. Input to Passivization

jari-Ø ñuca-ta cri-rca micu-naya-shca-ta
man-nom 1sg-acc believe-3 past eat-desid-past part-acc
'The man believes me to have wanted to eat.'

b. Output of Passivization
ñuca-ca jari-Ø cri-shca ca-ni
1sg-topic man-nom believe-past part be-1sg

micu-naya-shca-ta
eat-desid-past part-acc

'I was believed by the man to have wanted to eat.'

(26) a. Input to Passivization

jari-Ø can-da cri-n chiri-shca-ta
man-nom 2sg-acc believe-3 cold-past part-acc
'The man believes you to have been cold.'

(26) b. Output of Passivization

can-ga jari-Ø cri-shca ca-ngui
2sg-topic man-nom believe-past part be-2sg

chiri-shca-ta
cold-past part-acc

'You were believed by the man to have been cold.'

Thus, examples (23)-(26) indicate that accusative experiencers in both constructions are treated as subjects rather than as objects by SR.

The final rule to be discussed in this section is Passivization. Sentences involving this rule were given above in conjunction with our discussion of SR. What is relevant to the present argument is that Passivization promotes the non-subject to derived subjecthood and demotes the underlying subject to non-subject nominative. For example, in (27),

(27) ñuca-ca wawa-ta micu-chi-rca-ni
1sg-topic child-acc eat-caus-past-1sg
'I fed the child.'

the accusative noun phrase, wawa 'child', is the direct object. The verb agrees with ñuca 'I', the subject. In the output of Passivization as shown in (28),

(28) wawa-ca ñuca-Ø micu-chi-shca ca-rca
child-topic 1sg-nom eat-caus-past part be-3 past
'The child was fed by me.'

wawa 'child' has been promoted to subject, as is shown by third person verb agreement. As was noted previously, the promotion aspect of Passivization fails to apply to nominative subjects while the demotion aspect of the rule fails to apply to objects. Hence, an NP that is promoted by Passivization may be assumed not to be a subject prior to the application of the rule and an NP that is demoted by the rule may be assumed to be a subject in the input structure.

To turn to the two accusative experiencer constructions, the experiencer nominals fail to undergo promotion by Passivization, as is indicated for -naya- verbs in (29) and lexical experiencers in (30).

(29) a. Input to Passivization

ñuca-ta puñu-naya-rca
1sg-acc sleep-desid-3 past
'I would like to sleep.'

(29) b. Ungrammatical output of Passivization

*ñuca-ca puñu-naya-shca ca-rca-ni
lsg-topic sleep-desid-past part be-past-lsg
(I was slept.)

(30) a. Input to Passivization

ñuca-ta chiri-rca
lsg-acc cold-3 past
'I was cold.'

b. Ungrammatical output of Passivization

*ñuca-ca chiri-shca ca-rca-ni
lsg-topic cold-past part be-past-lsg
(I was colded.)

Thus, the experiencer nominals in both constructions behave like nominative subjects rather than like accusative objects with regard to promotion by Passivization.

The experiencer nominals also behave as subjects with regard to the demotion aspect of Passivization. This is shown for -naya- desideratives in (31) and lexical experiencers in (32).

(31) a. Input to Passivization

ñuca-ta-ca mishqui-ta micu-naya-rca
lsg-acc-topic candy-acc eat-desid-3 past
'I would like to eat candy.'

b. Grammatical output of Passivization of patient

mishqui-ca ñuca-Ø micu-naya-shca ca-rca
candy-topic lsg-nom eat-desid-past part be-3 past
'Candy was wanted to eat by me.'

c. Ungrammatical output of Passivization of experiencer

*ñuca-ca mishqui-{Ø / ta} micu-naya-shca ca-rca-(ni)
lsg-topic candy-{nom / acc} eat-desid-past part be-(3) past-(lsg)
(I was eaten by candy / I was eaten candy.)

(32) a. Input to Passivization

ñuca-ta-ca mishqui-ta muna-rca
lsg-acc-topic candy-acc want-3 past
'I wanted candy.'

(32) b. Grammatical output of Passivization of patient

mishqui-ca ñuca-Ø muna-shca ca-rca
candy-topic 1sg-nom want-past part be-3 past
'Candy was wanted by me.'

c. Ungrammatical output of Passivization of experiencer

*ñuca-ca mishqui-{Ø / ta} muna-shca ca-rca-(ni)
1sg-topic candy-{nom / acc} want-past part be-(3)past-(1sg)
(I was wanted by candy / I was wanted candy.)

The sentences of (31)-(32) show that the experiencer nominal undergoes demotion by Passivization as would be expected for subjects. This is in contrast with the patient, which is promoted to derived subject. Note that the patient may not undergo demotion nor the experiencer promotion.

These facts are predictable from the hypothesis that in the input to Passivization the experiencers are parsed as subjects and the patients as objects. But the counter-hypothesis that both accusative nominals are non-subjects would provide no explanation for the different behavior of the two groups of accusative nominals. Hence, we conclude that both aspects of Passivization provide evidence for the subjecthood of the experiencers in both constructions.

In this section we have examined the effect of three grammatical rules, Equi, SR, and Passivization, on the two experiencer constructions. With respect to each of these rules, the experiencers in both constructions behaved syntactically like subjects rather than like non-subjects. The rules did not distinguish between the constuctions in any way. In the section which follows, we shall see that this is not the case for all syntactic rules. For certain rules, the experiencer nominals in the desiderative construction behave more like nominative subjects than do experiencer nominals with lexical verbs of experience.

4. Switch reference constructions

In this section we shall consider the behavior of experiencer nominals in two types of switch reference constructions: adverbial clauses and purpose clauses. In these constructions, the suffixes used when the embedded subject is coreferential with that of the main clause differ from those used when the embedded and matrix subjects are not coreferential. This is illustrated for nominative subjects in the examples which follow.

(33) a. Adverbial clause: matrix and embedded subject coreferential

(ñuca-$Ø_i$) wasi-man chaya{-shpa / *-jpi} ñuca-$Ø_i$ can-da$_j$ ricu-ni
1sg-nom house-dat arrive{-coref adv / *-non-cor adv} 1sg-nom 2sg-acc see-1sg
'When I arrive home, I see you.'

(33) b. Adverbial clause: Matrix and embedded subjects non-coref.

ñuca-$\emptyset_i$ wasi-man chaya{-jpi / *shpa} can-$\emptyset_j$ ñuca-ta$_i$ ricu-ngui

lsg-nom house-dat arrive{-non-coref adv / -coref adverb} you-nom lsg-acc see-2sg

'When I arrive home, you see me.'

c. Purpose clause: matrix and embedded subjects coreferential

ñuca-$\emptyset_i$ wasi-man ri-ju-ni (ñuca-$\emptyset_i$)[12] can-da$_j$ ricu{-ngapaj / *-chun}

lsg-nom house-dat go-prog-lsg lsg-nom 2sg-acc see{-coref purp / *-non-cor purp}

'I am going home to see you.'

d. Purpose clause: matrix and embedded subjects non-coref.

ñuca-$\emptyset_i$ wasi-man ri-ju-ni can-$\emptyset_j$ ñuca-ta$_i$ ricu{-chun / *-ngapaj}

lsg-nom house-dat go-prog-lsg 2sg-nom lsg-acc see{-non-cor. purp / -coref. purp}

'I am going home for you to see me.'

As is shown in the examples of (33), the suffix -shpa is used in adverbial clauses in which the embedded and matrix subjects are coreferential. When these NPs are not coreferential, the suffix -jpi is found. In purpose clauses -ngapaj indicates coreference between matrix and embedded subjects. The suffix -chun shows non-coreference. It should be noted that when there is merely coreference between a non-subject nominal in the matrix clause and the embedded subject, -jpi and -chun are used, as was seen in (33b) and (33d). Similarly, -jpi and -chun are found when the matrix subject and an embedded non-subject nominal are coreferential:[13]

(34) Coreference between matrix subject and embedded non-subject

a. Adverbial clause

pai-$\emptyset_j$ ñuca-ta$_i$ ricu{-jpi / *-shpa} ñuca-$\emptyset_i$ pai-ta$_j$ maca-sha

3sg-nom lsg-acc see{-non-cor. adv / *-coref. adv.} lsg-nom 3sg-acc hit-lsg fut

'When he sees me, I will hit him.'

b. Purpose clause

ñuca-$\emptyset_i$ pai-ta$_j$ maca-rca-ni pai-$\emptyset_j$ waca{-chun / *-ngapaj}

lsg-nom 3sg-acc hit-past-lsg 3sg-nom cry {-non-coref purp / -coref purp}

'I hit him so he would cry.'

We shall now employ switch reference to determine whether accusative experiencers are treated as subjects or as non-subjects. It will be shown that the accusative experiencer nominals in -naya- desideratives behave syntactically like nominative subjects with respect to the assignment of adverbial and purposive suffixes. In contrast, accusative experiencers with lexical verbs of experience may be treated either as subjects or as non-subjects for purposes of adverbial and purposive suffix assignment.

We shall first consider -naya- experiencers. The sentences of (35) show that accusative nominals with -naya- experiencers are viewed as subjects for purposes of adverbial and purposive suffix assignment. This is shown by the use of -shpa and -ngapaj, which, as was seen above, are restricted to sentences in which the matrix and embedded subjects are coreferential. Note that the sentences of (35) would be ungrammatical if -jpi and -chun, the suffixes used when matrix and embedded subjects are non-coreferential, were used. It will be seen below that this is not the case for lexical experiencer nominals.

(35) Desiderative experiencers treated as subjects in switch reference clauses

a. Adverbial clause: desiderative matrix clause

(ñuca-$\emptyset_i$) chagra-pi trabaja{-shpa / *-jpi} ñuca-ta_i
lsg-nom field-in work {-coref. adv. / *-non-coref. adv.} lsg-acc

puñu-naya-n
sleep-desid-3

'When I work in the field, I want to sleep.'

b. Adverbial clause: desiderative embedded clause

(ñuca-ta_i) puñu-naya {-shpa / *-jpi} ñuca-$\emptyset$- ca_i puñu-ni
lsg-acc sleep-desid {-coref. adv. / *-non-coref. adv.} lsg-nom-topic sleep-1sg

'When I want to sleep, I sleep.'

c. Purpose clause: desiderative matrix clause

cunan tuta ñuca-ta_i puñu-naya-n caya maimi $\emptyset_i$
now night lsg-acc sleep-desid-3 tomorrow a lot

tragaja {-ngapaj / *-chun}
work {-coref. purp. / *-non-coref. purp.}

'I'd like to sleep tonight in order to work a lot tomorrow.'

(35) d. Purpose clause: desiderative embedded clause

ñuca-Ø-ca$_i$ ali micuna-ta rura-ni Ø$_i$
1sg-nom-topic good food-acc make-1sg

micu-naya {-ngapaj / *-chun}

eat-desid {-coref. purp. / *-non-coref. purp.}

'I make good food so that I will want to eat.'

Lexical experiencers exhibit a pattern different from that seen in (35). As is seen in (36), subject to certain variation from speaker to speaker to be discussed below, both sets of suffixes may be used: those indicative that the matrix and embedded subject are coreferential (-shpa and -ngapaj) and those indicative of non-coreferentiality of subjects (-jpi and -chun):

(36) Adverbial clauses with lexical experiencers treated as either subject or non-subject

a. Lexical experiencer in matrix clause

Ø$_i$ urcu-pi trabaja {-shpa / -jpi} jari-ta-ca$_i$

mountain-in work {-coref. adv. / -non-coref. adv.} man-acc-topic

yacu-ta muna-rca
water-acc want-3-past

'While working on the mountain, the man wanted water.'

b. Lexical experiencer in subordinate clause

ñuca-Ø$_i$ jambi-dur-man ri-rca-ni Ø$_i$ nana {-shpa / -jpi}

1sg-nom medicine-agt-dat go-past-1sg hurt {-coref. adv. / -non-coref. adv.}

'I went to the doctor hurting.'

Note that when the subject NPs are non-coreferential, as in (37), only -jpi is used:

(37) Adverbial clauses with lexical experiencer: non-coreferential subjects

a. Lexical experiencer in matrix clause

jari-Ø$_i$ urcu-pi trabaja {-jpi / *-shpa}

man-nom mountain-in work {-non-coref. adv. / *-coref. adv.}

(37) a. (continued)

wawa-ta-ca$_j$ yacu-ta muna-rca

child-acc-topic water-acc want-past

'While the man was working on the mountain, the child wanted water.'

b. Lexical experiencer in subordinate clause

ñuca-Ø$_i$ jambi-dur-man ri-rca-ni wawa-ta$_j$

1sg-nom medicine-agt-dat go-past-1sg child-acc

nana {-jpi / *-shpa}

hurt {-non-coref. adv. / *-coref. adv.}

'I went to the doctor when the child was hurting.'

With regard to purpose clauses, some variation was found among the speakers consulted. In the examples of (38), the symbol (#) indicates that the sentence in question was accepted by one informant and rejected by another.

(38) Purpose clauses

a. Lexical experiencer in matrix clause

ñuca-ta$_i$ trabaja-na-ta muna-n Ø$_i$ culqui-ta

1sg-acc work-infin-acc want-3 money-acc

japi {-ngapaj / #-chun}

take {-coref. purp. / #-non-coref. purp.}

'I want to work in order to make money.'

b. Lexical experiencer in subordinate clause

pai$_i$ larca-pi wambu-rca ama Ø$_i$ rupa {-ngapaj / #-chun}

3sg brook-in swim-3 past neg burn {-coref. purp. / #-non-coref. purp.}

'He swam in the water in order not to be hot.'

The examples of (38) show two patterns. For one speaker, switch reference in purpose clauses with lexical experiencers show the same pattern as in adverbial clauses: the experiencer NP may be analyzed either as a subject or as a non-subject. For the other speaker consulted, the lexical experiencer in a purpose clause could be treated only as a subject, a pattern like that found for -naya- experiencers. Parallel to what was found with adverbial clauses, when the lexical experiencer was not coreferential with the subject of the other clause,

-chun was invariably the only form accepted.

(39) Purpose clauses with lexical experiencers: non-coreferential subjects

a. Lexical experiencer in matrix clause

ñuca-ta-ca$_i$ taita-man$_j$ cara-na-ta muna-n
1sg-acc-topic father-dat servi-infin-acc want-3

sinlli $\emptyset_j$ ca{-chun / *-ngapaj}

strong be{-non-coref. purp. / *-coref. purp.}

'I want to serve it to father so he will be strong.'

b. Lexical experiencer in subordinate clause

ñuca-ca$_i$ wawa-ta$_j$ wambu-chi-rca-ni ama $\emptyset_j$
1sg-topic child-acc swim-caus.-past-1sg neg

rupa {-chun / *-ngapaj}

burn {-non-coref. adv. / *-coref. adv.}

'I caused the child to swim so he wouldn't be hot.'

We have shown in this section that -naya- desiderative experiencers and lexical experiencers differ in behavior with regard to switch reference constructions. Desiderative experiencers are invariably analyzed as subjects. Lexical experiencers, however, with some variation from speaker to speaker, may be analyzed either as subject or as non-subject with respect to switch reference. The synchronic and diachronic implications of these facts will be discussed in the following section. It is sufficient at present to note that a pattern similar to that found for CM and POC (section 2) is found with switch reference: desiderative experiencers are (in this case, obligatorily) treated as subjects while lexical experiencers are (in this case, optionally) treated as non-subjects.

5. Conclusions

In the preceding sections we examined the degree to which two types of accusative nominals, -naya- desiderative experiencers and lexical experiencers, exhibit morphological and syntactic properties otherwise restricted in IQ to nominative subjects. The properties associated with each type of experiencer are reviewed schematically in (40).

(40)	Subject Properties	Lexical Experiencers	Desiderative Experiencers
Behavioral Subject Properties	SR	yes	yes
	Passivization	yes	yes
	Equi	yes	yes
	Switch Reference Purpose	#	yes
	Switch Reference Adverbial	#	yes
Subject Coding Properties	PCC	no	yes
	CM	no	#
	VA	no	no

yes=behaves like nominative subject
no=does not behave like nominative subject
#=optional or individual variation

In (40) subject properties are divided into behavioral and coding properties. (Cf. footnote 5) Lexical experiencers exhibit some, though not all behavioral properties, while -naya- desideratives exhibit all behavioral properties, and, in addition some coding properties associated with IQ subjects. The question which naturally arises is whether there is an explanation for these facts. Does the pattern of subject properties seen in (40) reveal any significant diachronic or synchronic regularities in IQ?

The distribution of subject properties seen in (40) appears to have a principled explanation. It has been shown in previous work (Cole et al., 1978) that when a historically non-subject nominal comes to be a subject, properties associated with subjecthood are often acquired gradually rather than as a single historical change. Furthermore, in the process of historical change, behavioral subject properties are acquired prior to coding properties. The cross-linguistic patterns observed in Cole et al. (1978) suggest an analysis for the data described in this paper. The pattern seen in (40) is just what would be predicted on the basis of Cole et al. if both experiencer constructions are instances of grammatical change in progress. Viewed from this perspective, the distribution of subject properties would be taken as reflecting the degree of historical change of the experiencer from non-subject (presumably, object) to subject. But, although the accusative nominals in the two constructions are undergoing a similar change, the change would appear to be preceding independently in each construction. In the case of -naya- experiencers, the change would appear to be nearly complete. This is shown by the fact that both behavioral and coding properties are associated with -naya- experiencer nominals. In contrast, lexical experiencer nominals appear to be at an earlier stage in the same process of change.

The interpretation of (40) suggested by Cole et al., which constitutes, in effect, a partial internal reconstruction of the constructions in question, appears to be supported by the somewhat scant comparative data available. To the best of our knowledge, no detailed study has been made of the properties of experiencers in analogous constructions in other dialects of Quechua. Thus, we are not able to provide convincing arguments that experiencers lacked all behavioral subject properties in common Quechua. But it does seem fairly clear that the subject coding properties associated with -naya- experiencers are an innovation. In nearly all Quechua dialects -naya- experiencers are found in accusative case. The possibility of nominative is quite exceptional. Thus, in the absence of convincing counter-evidence, such as relevant archaisms, the most likely reconstruction would be the one in which the property found in most of the daughter languages, accusative case, is hypothesized to be a property of the proto-language. It would seem, then, that comparative evidence suggests that -naya- experiencers are innovative vis-a-vis lexical experiencers, the same conclusion drawn above on the basis of dialect internal evidence.

The synchronic analysis of the two constructions lends support to the hypothesis that the desiderative construction represents a more advanced stage in historical change than does the lexical experiencer construction. Harris (to appear,a) and Perlmutter (to appear) inter alia have proposed that experiencer constructions like those examined here involve a rule of Inversion. This rule demotes the underlying subject to object.[14] Such an analysis is well-motivated synchronically, because it provides as explicit explanation for the fact that the experiencer NP has both subject and non-subject properties. If an Inversion analysis is accepted, there are certain interesting consequences. The synchronic analysis of the two experiencer constructions, in conjunction with historical data presented by Harris on the development of experiencer constructions in other languages, provides additional evidence that the two experiencers are in transition from non-subject to subject, and that the -naya- experiencers have reached a more advanced stage than have the lexical experiencers.

An example of how Inversion allows the association of superficial object properties with the underlying subject is given in (42) which is roughly the derivation of (41):

(41) ñuca-ta puñu-naya-n
lsg-acc sleep-desid-3

(42) a. Underlying representation

ñuca (subject)	puñu-naya-
lsg	sleep-desid

b. Output of Inversion

ñuca (object)	puñu-naya-
lsg	sleep-desid

(42) c. Output of Case Marking

ñuca-ta puñu-naya-
1sg-acc sleep-desid

d. Output of Verb Agreement

ñuca-ta puñu-naya-n
1sg-acc sleep-desid-3

Once an Inversion analysis is assumed, certain rule orderings are necessary in order to generate the sentences found in the language.[15] According to this analysis, desiderative and lexical experiencer constructions differ only in terms of how Inversion is ordered with respect to other rules. The fact that desiderative experiencers are more subject-like than lexical experiencers is explained by ordering Inversion later in the derivation for -naya- than for lexical experiencers. The ordering of rules needed for each construction is presented in (43). We would like to emphasize that the rule ordering posited in (43) is motivated on purely synchronic grounds. The necessity for these orderings is justified briefly in footnote 15.

(43) Ordering of Inversion

a. Lexical experiencers

SR
Passivization
Equi
#Inversion
#Switch Reference Purpose
#Switch Reference Adverbial
POC
CM
VA

b. Desiderative experiencers

SR
Passivization
Equi
Switch Reference Adverbial
Switch Reference Prupose
POC
#Inversion
#CM
VA

#indicates variability in ordering

The rule ordering seen in (43) correctly predicts that -naya- experiencers will exhibit a wider range of subject properties than lexical experiencers. In general, the later Inversion is ordered vis-a-vis other rules, the more subject properties will be associated with the

experiencer nominal. Thus, the Inversion hypothesis, in conjunction with the rule orderings proposed in (43), accounts for the differences between the two constructions.

We shall turn now to the diachronic implications of the difference in rule ordering in (43a) and (43b). Harris (to appear, b) examines grammatical changes leading ultimately to the loss of non-nominative experiencer constructions in two unrelated languages, English and Udi. Employing an Inversion analysis, Harris considers the synchronic state of the languages at various stages in their development. She shows that in both cases Inversion was ordered prior to morphosyntactic coding rules (case marking and verb agreement) at an early stage in the development of the language, but that at later stages the rule was successively reordered so as to follow the coding rules. The effect of the reordering was to diminish, and eventually, to eliminate non-subect coding properties. Upon being reordered after all other rules, inversion ceased to have any surface effects, and the rule, presumably, was lost.

The changes reported for English and Udi provide cross-linguistic corroboration for the analysis ofIQ proposed above. We suggested that -naya- experiencers are at a later stage in the acquisition of subjecthood than are lexical experiencers. We shall assume that IQ would be expected to undergo a process of change similar to that found by Harris in English and Udi. On the basis of this assumption, our diachronic analysis predicts that Inversion should be ordered later in the derivation with -naya- experiencers than with lexical experiencers. This is, in fact, the rule ordering found in (43).[16]

To summarize our conclusions, -naya- desiderative experiencers manifest a considerably wider range of subject properties than do lexical experiencers. The differences between the two experiencer constructions suggest that, while experiencers in both constructions are in the midst of a change from objecthood to subjecthood, the change is more advanced in the desiderative experiencer construction than in the lexical experiencer construction.

NOTES and REFERENCES

[1]The dialect of Quechua described in this paper is spoken in northern Ecuador. There are between 40,000 and 60,000 speaker of IQ and between 6,000,000 and 8,000,000 speakers of Quechua as a whole. For a general description of IQ, see Stark (1973).

This paper is based on data from speakers born and raised in the town of Mariano Acosta, in the eastern part of the Province of Imbabura. We would like to express our great gratitude to Carmen Chuquín, our primary informant, now a graduate student in linguistics, with whom we worked extensively at the University of Illinois, and to Mariana and Emilia Chuquín, with whom we worked during a visit to Otavalo (Imbabura) in January 1978. Some aspects of the data reported on in this paper may be restricted to the subdialect spoken around Mariano Acosta.

We would like to thank Gaby Hermon, Wayne Harbert, and Chuck Kisseberth for their suggestions regarding the analysis of the data presented here and for comments on earlier versions of this paper. We would also like to thank Jean Hannah for finding a variety of inconsistencies and problems with the original manuscript. Her assistance in preparing the paper for publication was invaluable.

This work was supported in part by the National Science Foundation (SOC 7500244), the Research Board of the University of Illinois and the Center for International-Comparative Studies of the University of Illinois. The assistance of these institutions is gratefully acknowledged.

[2]Due to a phonological voicing rule, the accusative suffix is pronounced -da following nasals and -ta elsewhere.

[3]The verb muna- 'want'may also be used in 'personal' constructions with a nominative experiencer, in which case it agrees with the nominative NP in person and number.

[4]In addition, -naya- may be added to a limited number of nouns, e.g. yacu 'water', to form a desiderative verb, yacu-naya- 'desire to drink, be thirsty'.

Both lexical and desiderative verbs of experience may be converted into 'personal' verbs, taking nominative experiencer. This is done by the addition of the morpheme -chi: ñuca-ta puñu-naya-n (experiencer in accusative, third person agreeement) but ñuca puñu-naya-chi-ni (experiencer in nominative, first person singular agreement). Both sentences are glossed identically as 'I want to sleep' by speakers. The suffix -chi is normally a causative morpheme, but that does not appear to be its function here. We shall have no further discussion of this construction.

It should be noted that in -naya- constructions the experiencer is understood to be identical to the subject of the verb stem. That is, the construction may be used to express the notion 'I would like to eat', but not 'I would like Harry to eat'.

[5]Following Keenan (1976), we distinguish between subject coding properties and subject behavioral properties. The former includes rules involving the morphosyntactic coding of grammatical relations, such as case marking and verb agreement. The latter includes non-morphological syntactic rules, such as deletion under identity and subject raising.

For a justification of the distinction, see Keenan (1976) and Cole et. al. (1978). The distinction between coding and behavioral properties plays a major role in section 5.

[6]Passivization in Imbabura promotes an object NP to subject, demotes transitive subjects to non-subject nominatives and makes appropriate changes in verbal morphology. The two aspects of the rule, promotion and demotion, constitute tests for objecthood and subjecthood:

Objects, both direct and indirect, though not obliques, may be promoted to subject, but subjects may not. Subjects are demoted to non-subject nominatives by Passivization.

[7]TOC in IQ, and in Ecuadorian dialects generally, is restricted to first person objects. In Peruvian dialects, second person objects may be cliticized as well. The rule is optional in IQ. A full NP object may co-occur with the clitic for emphasis.

[8]When experiencers with -naya- verbs appear in the nominative, they are invariably marked as topics (-ca). Topic marking is not restricted to any single type of NP such as subjects. It is frequently used to disambiguate subject from non-subject, as in (i):

(i) ñuca-ta ??-Ø / -ca pai-ta ricu-na-ta muna-n

1sg-acc-topic 3sg-acc see-infin.-acc want-3

'I want to see him.'

The potential ambiguity in the version of (i) without -ca arises from the fact that NPs in Quechua are frequently conjoined by juxtaposition. Thus, the version of (i) without -ca is, normally, interpreted as a conjoined object construction (unspecified subject). The use of -ca shows that ñuca-ta and pai-ta are not a single constituent. The topic marker has a similar function in the passive, as is shown below.

[9]SR is governed by approximately the same class of verbs that governs subject to object raising in English: verbs of perception (e.g., ricu- 'see', uya- 'hear') and verbs of propositional attitude (e.g., cri- 'believe', yacha- 'know').

[10]Complement clauses and, in fact, nearly all subordinate clauses in IQ are non-finite. The verb of the complement clause receives a nominalizing suffix which indicates relative time reference vis-a-vis the matrix clause: The suffix -shca indicates past time reference, -j and -y indicate present and -ng indicates future. The same suffixes (with the exception of -y) are found in relative clauses. In addition, -shca has a role analogous to that of past participle in the formation of the perfect and of the passive.

[11]TOC is not always a clear test for SR. With certain matrix clause verbs (e.g., muna- 'want') clitic climbing takes place: A complement object cliticizes to the matrix verb, as in (i):

(i) pai-Ø muna-wa-n ayuda-na-ta

3-sg-nom want-1-3 help-infin-acc

'He wants to help me.'

The verb cri- 'believe' is employed in (22) because clitic climbing is not possible with cri-, as is shown by the ungrammaticality of (22b). For further discussion of clitic climbing, see Muysken (1977).

[12]The retention of the optional noun phrase is, in this example,

and in several others in this section, grammatical but stylistically objectionable.

[13]We were unable to obtain clear informant judgements regarding whether the assignment of the suffixes in question is strictly controlled by derived grammatical relations. Our informant failed to have consistent intuitions with respect to the suffixes needed in adverbial and purpose clauses subordinated to passive matrix clauses.

[14]According to Perlmutter, and Harris, the rule universally demotes a subject to indirect object. The appearance of the underlying subject as a direct object in IQ (as evidenced by accusative case) would presumably be accounted for by the subsequent promotion of the indirect object to direct object. We shall take no position on whether such a derivation is in fact justified in IQ.

[15]The examples presented previously provide the basis for the rule ordering statements which follow. It is assumed that rule ordering is transitive: If A precedes B and B precedes C, then A precedes C.

SR feeds Passivization, and thus, must precede it. Equi must precede Inversion or Inversion would bleed Equi. With lexical experiencers, Inversion may precede or follow both switch reference rules for some speakers. For others it must follow Switch Reference Purpose and precede Switch Reference Adverbial. Thus, Switch Reference Purpose precedes Switch Reference Adverbial. (see section 4 for details.) In some derivations, Inversion follows the switch reference rules but precedes POC. Thus, the switch reference rules must precede POC.

With desiderative experiencers, Inversion may come between POC and CM. Hence, POC precedes CM. It may also follow CM but precede Verb Agreement. Therefore, Verb Agreement must follow CM.

[16]The historical acquisition of subject properties does not always involve the reordering of a rule in the derivation. See Cole et al. for details.

REFERENCES

Cole, Peter, Wayne Harbert, Gabriella Hermon, and S.N. Sridhar. 1978. On the acquisition of subjecthood. Studies in the Linguistic Sciences 7.2

Harris, Alice. To appear, a. Inversion as a rule in universal grammar: Georgian evidence. Studies in Relational Grammar, ed. by David Perlmutter.

_____ To appear, b. On the mechanism of syntactic rule loss: evidence from English and Udi. Studies in Relational Grammar, ed. by David Perlmutter.

Keenan, Edward. 1976. Towards a universal definition of 'subject'. Subject and Topic, ed. by Charles Li. Academic Press.

Muysken. 1977. Clitics as Moving Features. Manuscript.

Perlmutter, David. To appear. Evidence for inversion in Russian, Japanese, and Kannada. Studies in Relational Grammar,ed. by David Perlmutter.

Stark, Louisa and Lawrence Carpenter. 1973. El Quichua de Imbabura: una gramatica pedagogica. Instituto Interandino de Desarrollo. Otavalo, Ecuador.

Studies in the Linguistic Sciences
Volume 8, Number 1, Spring 1978

WHY DYIRBAL ISN'T ERGATIVE AT ALL*

Janice Jake

The recent interest in ergativity has underscored the singular position of Dyirbal as a syntactically ergative language. Even among the morphologically ergative aborigine languages of Australia, Dyirbal represents a unique situation.[1] In most morphologically ergative languages it can be easily demonstrated that the morphology does not directly represent the grammatical relations reflected by the syntactic rules of the language. Dixon (1972), however, has claimed that absolutive and ergative, rather than subject and object, are the relevant grammatical relations in Dyirbal. This paper examines the evidence supporting Dixon's claim that Dyirbal is syntactically ergative and shows how this evidence is equally supportive of a nominative-accusative analysis in which subject and object are the grammatical relations referred to by the syntactic rules of Dyirbal. In addition to the data provided by Subject-Subject-Equi-NP-Deletion, Conjunction Reduction, Relativization and Reflexivization discussed in Dixon (1972), Imperative Deletion and Object-Subject-Equi-NP-Deletion offer two sets of data which can adequately be accounted for by a subject-object analysis, but not by an absolutive-ergative analysis.

Prior to an examination of Dyirbal, a few examples from Tongan and Avar are presented to show the typical situation in morphologically ergative languages. Then Dixon's analysis of Dyirbal as a syntactically ergative language will be compared to an alternative subject-object analysis. The two analyses are then evaluated in terms of cross-linguistic evidence and the predictions they make for the other parts of the grammar of Dyirbal. Here not only do cross-linguistic precedence arguments favor the subject-object analysis, but the language internal evidence, comprised of two sets of data which the subject-object analysis can adequately account for, but which the absolutive-ergative analysis cannot, includes a counterexample to Dixon's analysis.

Ergativity has traditionally referred to languages in which intransitive subjects and direct objects ate marked by the same morphological form and transitive subjects are marked by another morphological form. In some languages morphological ergativity is indicated by case marking, as shown by the Tongan examples in (1):

(1) Tongan (Chung, 1976a)

a. Ergative Case Marking:

Verb	Ø or 'a Subject	(intrans)
e Subject	Ø or 'a Direct Object	(trans)

b. (intrans) 'oku fu'f hela'ia 'a 'Ālani 'i he ŋgāue
prog very tired abs Alan caus the work
'Alan is very tired because of work.'

c. 'oku fa'u 'e hoku tokuoa 'a e tēpile
prog make erg my sibling abs the table
'My brother is making a table.'

Verb agreement can also be ergative, as it is in Avar:

(2) Avar (Comrie, forthcoming)
a. vas v- ekerule
boy-abs sg/masc/abs-run
'The boy runs.'
b. jas j- ekerule
girl-abs sg/fem/abs-run
'The girl runs.'
c. vas-as: jas j- ec:ula
boy-erg girl-abs sg/fem/abs-praise
'The boy praises the girl.'

In morphologically ergative languages the NPs which are in nominative case, as in (1), and the NPs which control verb agreement, as in (2), have been referred to as absolutives. Transitive subjects are referred to as ergatives.

Since nominative case marking and control of verb agreement have been generally regarded as properties of subjects, the applicability of the notion subject has been questioned for ergative languages. A more detailed examination of subjects shows that in addition to morphological properties like nominative case marking and control of verb agreement being cross-linguistically characteristic of subjects, there are also certain syntactic properties associated with subjects.[2] Some of the properties are the ability to delete under Equi-NP-Deletion and Conjunction Reduction, the ability to control but not undergo clause-bounded Reflexivization, and the ability to undergo Subject-to-Subject Raising. In morphologically ergative languages in a transitive sentence, it is not the nominative NP which syntactically patterns with the intransitive subject, but the ergative NP. The examples in (3) show that Equi-NP-Deletion deletes subjects rather than absolutives in Tongan.

(3) a. pea na'e 'alu 'a e tangatá 'o ʰaa'i 'a e kulī
and past go abs the man comp sail far
'Then the man went andsailed away.'
b. na'e 'alu 'a e tangatá 'o ʰaa'i 'a e kulī
past go abs the men comp hit abs the dog
'The man went and hit the dog.'
c. *na'e 'alu 'a e fefine 'o u'u 'e he kulī
past go abs the woman comp bite erg the dog
(The woman$_i$ went the dog bit $\emptyset_i$)

In (4) both transitive and intransitive subjects, but not objects, undergo Subject-to-Subject-Raising in Tongan:

(4) a. 'e lava 'a e pēpē 'o lea
unm can abs the baby comp talk
'The baby can talk.'

b. 'oku ou lava 'o ongo'i 'a e tāme'a mei he fale ko ē
prog I can comp hear abs the music from the house pred that
'I can hear the music (coming) from that house.'

c. *na'e lava 'a e fefine 'o taa'i 'e Sione
past can abs the woman comp hit erg Sione
(The woman$_i$ can John hit $\emptyset_i$.)

As exemplified by Tongan, morphological ergativity does not correspond to syntactic ergativity. The grammatical relation of subject is applicable to morphologically ergative languages. The term absolutive only reflects the common morphology of the intransitive subject and the direct object. Ergative is not a grammatical relation but a case marking used for transitive subjects.

In order for a language to be syntactically ergative the syntactic rules of the language must treat absolutive and ergative as grammatical relations. Johnson (1974) attempted to characterize a hierchary of grammatical relations which would be applicable to syntactically absolutive-ergative languages as well as subject-object languages.[3] Grammatical relations were ranked, the specific relation of each rank being determined by the particular grammatical relations a language possessed. Thus absolutive represents the highest grammatical relation in Dyirbal, parallel to the relation of subject in languages like English. The ergative is the equivalent of the second highest relation of direct object. In Dyirbal the dative case marks an oblique or chômer similar to the passive agent by-phrase in English. In the following discussion of Dyirbal this terminology will be adapted to facilitate a comparison of the analyses. Absolutive, ergative, oblique, subject, and direct object will refer to grammatical relations; nominative, instrumental, and dative will be reserved for case marking.[4]

The subject of an intransitive sentence in Dyirbal is in nominative, as shown in (5). The sentences in (6) show the three possibilities for expressing an (at least underlyingly) transitive sentence. All three sentences fundamentally express the same thing: in (6a) the relationship between patient, agent, and verb is viewed as one event; in (6b) the dative NP is affected by the action of the agent; in (6c) the degree of success of the action on the patient is not clear.[5]

(5) **bayi yaṛa bani-ɲu**
cm n man n come-non-fut
'man came'

(6) a. balan ɟugumbil baŋgul yaṛaŋgu balgan
cm n woman n cm i man i hit-non-fut

b. bayi yaṛa bagun ɟugumbilgu balgal-ɲa-ɲu
cm n man n cm d woman d hit-ɲay-non-fut

c. bayi yaṛi baŋgun ɟugumbiṛu balgal-ɲa-ɲu
cm n man n cm i woman i hit-ɲay-non-fut
'man hit woman'

In an analysis treating Dyirbal as an absolutive-ergative language, (6a) would be considered to be closest to the underlying structure of a transitive sentence. The transitive subject is in instrumental case, which corresponds to the ergative relation and the object is in nominative case, as is the intransitive subject in (5). The other two sentence types could

be transformationally derived from the structure underlying (6a). In (7) Antipassive promotes the underlying ergative to absolutive and the underlying absolutive is demoted to an oblique NP.[6]

(7) a. balan ɟugumbil baŋgul yaraŋgu balan
cm n woman n cm i man i hit-non-fut ⟹
(absolutive) (ergative)
b. bayi yara bagun ɟugumbilgu balgal-ŋa-ŋu
cm n man n cm d woman d hit-ŋay-non-fut
(absolutive) (oblique)

After Antipassive has applied to promote the underlying ergative and demote the underlying absolutive, a rule of Dative Promotion could apply to promote the oblique NP to an ergative relation as shown in (8).[7]

(8) a. bayi yara bagun ɟugumbilgu balgal-ŋa-ŋu
cm n man n cm d woman d hit-ŋay-non-fut ⟹
(absolutive) (oblique)
b. bayi yara baŋgun ɟugumbiru balgal-ŋa-ŋu
cm n man n cm i woman i hit-ŋay-non-fut
(absolutive) (ergative)

This absolutive-ergative analysis of Dyirbal accounts for the fact that only absolutives can be deleted by Equi-NP-Deletion and Conjunction Reduction and that only absolutives can undergo Relativization and trigger but not undergo Reflexivization, as shown in the following examples.

Equi-NP-Deletion: only nominative NPs delete
(9) balan ɟugumbil baŋgul yaraŋgu manda-n
cm n woman n cm i man i take-non-fut
'man took woman'
(10) a. balam miran baŋgun ɟugumbiru babi-n
cm n beans n cm i woman i scrape-non-fut
b. balan ɟugumbil bagum miraŋgu babil-ŋa-ŋu
cm n woman n cm d beans d scrape-ŋa-non-fut
'woman scrape beans'
(11) a. *balan ɟugumbil baŋgul yaraŋgu munda-n balam miran babi-n
cm n woman n cm i man i take-non-fut cm n beans-n scrape-non-fu
b. balan ɟugumbil baŋgul yaraŋgu munda-n bagum miraŋgu babil-ŋa-ŋu
cm n woman n cm i man i take-non-fut cm d beans d scrape-ŋay-non
'man took woman to scrape beans' fut

The ungrammaticality of (11a) shows that the instrumental NP in (10) cannot be deleted under identity with the nominative one in (9), although the nominative NP in (10b) can be deleted, as shown in (11b).

Conjunction Reduction: only nominative NPs delete
(12) bayi yara bani-ŋu
cm n man n come-non-fut
'man came'
(13) a. balan ɟugumbil baŋgul yaraŋgu balga-n
cm n woman n cm i man i hit-non-fut
b. bayi yara bagun ɟugumbilgu balgal-ŋa-ŋu
cm n man n cm d woman d hit-ŋay-non-fut
'man hit woman'

(14) a. *bayi yaṛa bani-ɲu balan ɖugumbil balga-n
cm n man n come-non-fut cm n woman n hit-non-fut
b. bayi yaṛa bani-ɲu bagun ɖugumbilgu balgal-ŋa-ɲu
cm n man n come-non-fut cm d woman d hit-ŋay-non-fut
'man came and hit woman'

The nominative NP in (12) can delete the nominative NP in (13b) under identity but not the instrumental NP in (13a).

Relativization: only nominative NP can be relativized
(15) a. bayi yaṛa bagal-ŋa-ɲu bagul yuṛigu banagaɲu
cm n man n /spear-ŋay-rel cm d kangaroo d/ return-non-fut
'man who speared kangaroo is returning' (bayi yaṛa ⟹ Ø)
b. *bayi yaṛa /baga-ɲu bayi yuṛi/ banagaɲu
cm n man n spear-rel cm n kangaroo n return-non-fut
(man who speared kangaroo is returning) (bagul yaṛagu ⟹ Ø)
c. *bayi yaṛa /balan ɖugumbil bura-ŋa-ɲu/ wayndin
cm n man n cm n woman n watch-ŋay-rel go uphill-non-fut
(man watched by woman is going uphill) (bagul yaṛagu ⟹ Ø)

Reflexivization: only nominative NP can control clausebounded Reflexivization
(16) a. bayi yaṛa buyba-yiri-ɲu
cm n man n hide-reflex-non-fut
b. *bagul yaṛagu buyba-yiri-ɲu
cm i man i hide-reflex-non-fut
c. *bagul yaṛagu buyba-yiri-ɲu
cm d man d hide-reflex-non-fut
'man hid himself'

This absolutive-ergative analysis will be compared with the following subject-object analysis. In this analysis (6a) is claimed to be closest to the underlying structure of transitive sentences.[8] The subject is marked with nominative case and the object with dative case. The -ŋay suffix marks the verb as being active transitive. In the derivation of (6a) Passive applies, promoting the object to subject and demoting the subject to an oblique NP marked by instrumental case. The verb is intransitivized by the deletion of -ŋay. Object Demotion would apply in the derivation of sentences like (6c). The object is demoted to an oblique NP and marked by the instrumental case.

The subject-object analysis also accounts for the facts about Equi-NP-Deletion, Conjunction Reduction, Relativization, and Reflexivization in (9)-(16). The nominative NPs which were absolutives in the absolutive-ergative analysis are subjects in the subject-object analysis. Both analyses work equally well in accounting for these data because they are mirror images of each other. Each analysis chooses one sentence whose structure is closest to the underlying form of a transitive sentence and derives the other two sentence types by appropriate transformations. What is underlying in one analysis is derived in the other and vice versa. Because the nominative case reflects the highest grammatival relation in both analyses, transformations referring to surface subjects or surface absolutives will not distinguish between the two analyses. One way to evaluate the competing analyses is to find other areas of the grammar of Dyirbal for which the two analyses make different predictions and see which one is correct. Rules

referring to the second highest grammatical relation, such as Object-Subject-Equi, ought to, as well as rules referring to underlying subjects.[9] Two sets of data dealing with such areas of the grammar of Dyirbal are discussed below.

Lower subjects can be deleted only under identity with terms. Thus, in English a direct object can trigger deletion, but a passive agent (chômer) cannot.

(17) a. Mary told Bob to go.
b. *Mary was told by Bob to go.
(with the reading 'Bob go')

In Dyirbal a dative NP can trigger deletion of a nominative NP, but an instrumental NP cannot.

(18) a. bayi yaṛa bagun ḍugumbilgu giga-ṇa-ṇu bagum miraŋgu
cm n man n cm d woman d tell-ŋay-non-fut cm d beans d
babil-ṇa-ṇu
scrape-ŋay-non-fut
'man told woman to scrape the beans'
b. *bayi yaṛa baŋgun ḍugumbiru giga-ṇa-ṇu bagum miraŋgu
cm n man n cm i woman i tell-ŋay-non-fut cm d beans d
babil-ṇa-ṇu
scrape-ŋay-nu
(man told woman to scrape beans)
c. *bayi yaṛa bangun ḍugumbiru giga-n bagum miraŋgu
cm n man n cm i woman i tell-non-fut cm d beans d
babil-ŋay-ṇu
scrape-ŋay-non-fut
(woman$_i$ told man (for) $\emptyset_i$ to scrape beans)

A subject-object analysis correctly predicts that objects (in dative case) can trigger deletion, but that oblique (instrumental) NPs cannot. An absolutive-ergative analysis incorrectly predicts that ergatives (instrumental NPs) should be able to trigger deletion and that dative case NPs should not be able to. Thus, the subject-object analysis correctly predicts that dative NPs represent the direct object grammatical relation and that instrumental NPs are not terms.

Another area of the grammar which provides evidence distinguishing between the two analyses is Imperative Deletion.[10] The following three sentences demonstrate how Imperative Deletion applies in Dyirbal.

(19) a. (ninda) bani
(2sg-n) come
'come!'
b. (ninda) bagul yaṛagu balgal-ṇa
(2sg-n) cm d man d hit-ŋay
'hit the man!'
c. (ninda) bayi yaṛa balga
(2sg-n) cm n man n hit
'hit the man!'
($\left\{\begin{matrix}\text{let}\\ \text{may}\end{matrix}\right\}$ the man be hit!)

A subject-object analysis predicts that subjects of imperatives will be deleted and an absolutive-ergative analysis predicts that absolutives will be deleted. An examination of the examples in (19) shows that only in the subject-object analysis is it possible to capture the relevant generalization about Imperative Deletion. In (19a) the subject of an active intransitive sentence is deleted in the subject-object analysis. In the absolutive-ergative analysis the absolutive is deleted. In (19b) the transitive subject of an active transitive sentence is deleted in the subject-object analysis. In the absolutive-ergative analysis a derived absolutive (underlying ergative) is deleted from a derived intransitive sentence. In (19c) the passive agent (or underlying subject) of a derived intransitive sentence is deleted. In the absolutive-ergative analysis the ergative is deleted from an active transitive sentence.

For the subject-object analysis the generalization about Imperative Deletion is that underlying subjects delete. For the absolutive-ergative analysis there is no clear generalization referring to only one grammatical relation. It appears that in intransitive (derived and underlying) sentences, the absolutive is deleted and that in transitive sentences, the ergative is deleted. Evaluation of the two generalizations shows the subject-object analysis to be preferred. Not only does it capture the generalization that the underlying subject deletes, but is does so without resorting to unprecedented grammatical relations.

Perhaps one objection to the subject-object analysis of Imperative Deletion in Dyirbal is that it allows passive imperatives. This objection is easily answered by the fact that other languages, like Maori, Javanese, and Bahasa Indonesian, also allow passive imperatives. In Maori the distribution of the active sentence is so marked that active imperatives are ungrammatical, as shown in (20). Imperative Deletion applies to delete the passive agent, that is, the underlying subject.

(20) Maori (Chung, 1976a)
  a. tua-ina te rakau
     fell-pass the tree
     'Fell the tree!'
  b. *tua-i te rakau
     fell-act the tree
     (Fell the tree!)
  c. puhi-a
     shoot-pass
     'Shoot (it)!'
  d. *patu te pooka
     hit-act the pig
     (Kill the pig!)

In passive imperatives in Javanese the underlying subject is also deleted, as shown in (21).

(21) Javanese (Horne, 1961)
  a. Môntóré di-adepké rônô.
     car pass-head over that way
     'Head the car over that way.'
  b. Kertasé di-tulisi djenengmu!
     paper pass-write name-your
     'Write your name on the paper!'

Thus Dyirbal is like other languages in which the passive is relatively unmarked. In these languages the generalizations about Imperative Deletion seems to be to delete the underlying subject. The subject-object analysis can correctly account for the facts about Imperative Deletion in Dyirbal, since Dyirbal behaves like other subject-object languages in which the distribution of the passive construction is less restricted than in English. This cross-linguistic data argues against Dixon's (1977) position that passives are universally the more morphologically complex and restricted in distribution than the active. There are languages in which the distribution of passive is less restricted than that of the active. The situation in Maori (Chung, 1976a) parallels that found in Dyirbal. The passive occurs in main clauses and the active usually occurs in reduced or subordinate clauses. While all main clauses in Maori are passive, some main clause in Dyirbal do occur in the active, as in the topic chain in (22).

(22) balagara ɟurmangu buralŋanu
two people-n track-d see-ŋay-non-fut
'the two (women) saw the (man's) track (i.e. the path worn by his constant travel to and from the scrub).' (Text XV, 12)

In languages like Javanese and Bahasa Indonesian the passive also has a relatively unrestricted distribution. Thus, when one is faced with deciding between a subject-object analysis with an unmarked passive similar to passive in other languages or an absolutive-ergative analysis which proposes grammatical relations which have not been attested in any other language, the preferred analysis is the subject-object analysis. Although the syntactic evidence from Imperative Deletion and Object-Subject-Equi indicates the active transitive sentence is the more morphologically complex -ŋay construction, Dixon maintains that this construction is derived and that Dyirbal is syntactically absolutive-ergative. This position entails the necessity of such a language being able to refer to more than one set of primary grammatical relations. This complication of the grammar of Dyirbal is avoided in the subject-object analysis.[11]

There are two possible objections which might be raised against the subject-object analysis of Dyirbal based on morphology: first, the active transitive form of the verb is morphologically more complex than the intransitive form, and, second, while the -ŋay suffix is absent in the Passive derived sentence, it is present in the Object-Demotion derived sentence.

In answer to the objection that the active verb morphology is more complex, it can be pointed out that in the Malaysian languages in which there is more than one kind of passive, the non-intentional passive is less complex than the active transitive.

(23) Bahasa Indonesian (R. MacDonald, 1976)
a. Ali men-yangka saya orang Palembang
Ali trans-thought I man Palembang
'Ali thought that I was a Palembang man.'
b. Saya Ali sangka orang Palembang
I Ali thought man Palembang
'Ali thought that I was a Palembang man' (passive)
c. Ali meng-anggap soal itu beres
Ali trans-considered problem the settled
'Ali considered the problem settled.'

(23) (con't)
d. Soal itu Ali anggap beres
problem the Ali considered settled
'Ali considered the pboblem settled.' (passive)

Thus it is not the case that crosslinguistically the passive form of the verb is more complex than the active.

With regards to the objection that the -ŋay suffix is not deleted in Object Demotion derived sentences, two replies are available. First, it is not the case that morphology always corresponds directly to the syntax of a language. For example, in Walbiri, a morphologically ergative language, the subject of an Antipassive derived sentence remains ergatively marked.

(24) Walbiri (Comrie, forthcoming)
a. Nat^{y}ululu ka-ŋa wawiri ḻuwaṇi
1sg-erg tense-1sg kangaroo-abs shoot
'I am shooting the kangaroo.'
b. Nat^{y}ululu ka-na-la-t^{y}inta wawiri-ki ḻuwani
1sg-erg tense-1sg-3sg-dat-intran kangaroo-dat shoot
'I am shooting at the kangaroo.'

Comrie(forthcoming) points out that in Basque the subject remains ergative and the auxillary verb is marked with a third person singular patient after Object Deletion applies, as shown in (25). Catford (1975) discusses ergative marking on subjects in Chechen and Dargi after Object Incorporation or Object Deletion, rules which are similar to an Object Demotion Antipassive.

(25) Basque (Bakaikoa, personal communication)
a. guk zuei artikulua irakurri d-i-zue-gu
1pl 2pl-dat article-def read 3sg-aux-2pl-1pl
'We read the article to you.'
b. guk zuei irakurri d-i-zue-gu
1pl 2pl-dat read 3sg-aux-2pl-1pl
'We read to you.'

Perhaps a second reason why the -ŋay suffix is not deleted in the Object Demotion construction in Dyirbal is the ambiguity which would result from the merger of the Passive with the Object Demotion derived structure. Both have nominative and instrumental NPs. The free word order in Dyirbal would not distinguish whether the nominative NP was an underlying subject and Object Demotion had applied or whether it was a derived subject and Passive had applied. The presence or absence of -ŋay distinguishes between these two constructions involving a nominative NP and an instrumental NP.

One problematic area for both analyses of Dyirbal is the pronoun system. The chart in (27), taken from Dixon (1972), has been annotated to make explicit the correspondences of case marking to the grammatical relations of each analysis. In the absolutive-ergative analysis absolutive intransitive subjects and absolutive objects have different case marking. Transitive subjects are marked with the same form as intransitive subjects (except for singular pronouns in the Giramay dialect). Dative case represents absolutive objects. In the subject-object analysis tha dative case represents the direct object relation. Dixon's (O) forms are subjects derived from Passive. The (S) and (A) forms represent underlying subjects; passive agents have the same

pronoun form as do other underlying subjects, that is, the transitive subjects in the -ŋay construction and underlying intransitive subjects. A passive subject is distinguished form surface subjects which were underlying by the suffixation of -ŋa of -ɲa to the pronoun form referring to underlying subjects.

This use of -ŋa/-ɲa is also found optionally with proper and common names, usually referring to humans and other animates. Thus object NPs (in dative case) have two forms; so do underlying objects which have been passivized to derived subjects.

(26) underlying subject: búrbula (only) 'búrbula' (name)
underlying object: búrbulagu, búrbalaɲangu
derived passive subject: búrbula, búrbalaɲa

This optional process seems to have been made obligatory for derived pronominal subjects. The problem of mapping the case morphology of the pronoun system onto grammatical relations has a simple solution in the subject-object analysis: underlying subjects have one form; derived passive subjects have an obligatory suffix; surface objects have another form. In the absolutive-ergative analysis there does not seem to be a clear generalization about the relationship of case marking in the pronoun system to the grammatical relations of absolutive and ergative; one form is used for all derived subject absolutives, all ergatives, and some underlying subject absolutives. These data have led Dixon (1977) to conclude that some parts of languages can and/or must refer to subjects.[12]

(27) Dyirbal pronoun system (Dixon 1972:50)

Abs-Erg Analysis	Sintrans/Atrans (G:Sonly)	(G:Aonly)	abs object	obliqueNP dative	person-number
Sub-Obj Analysis	underlying Sub (G:Sintrans)	(G:Strans)	derived passive sub	object NP	person-number
G	ŋayba	ŋaɖa	ŋaɲa	ŋaygungu	1sg
DM	ŋaɖa		ŋayguna	ŋaygungu	
G	ŋali		ŋaliɲa	ŋaliɲangu	1du
D	ŋaliɖi		ŋaliɖina	ŋaliɖingu	
M	ŋali		ŋalina	ŋalingu	
G	ŋana		ŋanaɲa	ŋanaɲangu	1pl
D	ŋanɖi		ŋanaɖina	ŋanaɖingu	
M	ŋana		ŋanana	ŋanangu	
G	ɲinba	ɲinda	ɲína	ɲinungu	2sg
DM	ɲinda		ɲinuna	ɲinungu	
G	ɲubilaɖi		ɲubilaɖina	ɲubilaɖingu	2du
D	ɲubalaɖi		ɲubalaɖina	ɲubalaɖingu	
M	ɲubala		ɲubalana	ɲubalangu	
G	ɲura		ɲuraɲa	ɲuraɲangu	2pl
D	ɲuraɖi		ɲuraɖina	ɲuraɖingu	
M	ɲuray		ɲurana	ɲurangu	

In summary, it seems that all of the evidence presented in this paper leads to the conclusion that Dyirbal is not syntactically ergative. A comparison of the two competing analyses has shown the subject-object analysis not only to have crosslinguistic precedence, but also to make correct predictionc about specific facts about Dyirbal for other parts of the grammar. It was first pointed out that while the subject-object analysis made the implicit claim that the passive was unmarked in Dyirbal, this situation was found to occur in other languages, Maori representing an extreme case, in which the active has a very marked distribution. The fact that such languages may not be as common as those in which the active is unmarked does not justify positing the elsewhere unattested grammatical relations of absolutive and ergative.

The subject-object analysis was shown to make correct predictions about the grammatical relation of dative NPs. The were shown to be able to trigger Object-Subject-Equi-NP-Deletion, an argument for their status as the second highest grammatical relation. In Imperative Deletion the subject-object analysis was able to capture the generalization that underlying subjects were deleted. It was shown that in other languages allowing passive imperatives, the passive agent, that is, the underlying subject, is also deleted. Maori provided an example of a language in which imperatives could only be passive.

Two possible objections about the verbal morphology in the subject-object analysis were shown to have parallels in other languages. The relative complexity of the transitive sentence is not unprecedented. The retention of the -ŋay suffix in sentences with demoted objects was shown to have parallels in the retention of ergative case marking in formally intransitive sentences in other languages, again demonstrating that surface morphology is not always a valid indication of the the syntax of a language. A disambiguating function was also mentioned. The pronoun system of Dyirbal was presented and shown to have a relationship with the grammatical relations of subject and object, but not absolutive and ergative.

While not conclusive, I hope that this initial evaltution of these two competing analyses has shown that not only is the subject-object analysis to be preferred because it accounts for the core data without recourse to unprecedented grammatical relations, but because crosslinguistic evidence and language internal facts show that of the two analyses, only it can adequately account for Dyirbal syntax.

NOTES

*Research for this paper has been supported by NSF grant SOC75-00244 and the Research Board, University of Illinois. I would like to thank C.C. Cheng, Peter Cole, Wayne Harbert, Gabi Hermon, Kim Hodges, Chuck Kisseberth, David Odden, and Anne Roalef for their help and comments. Earlier versions of this paper were presented the University of Wisconsin (Madison) and the University of Texas (Austin). Comments received at these times have been helpful and are gratefully acknowledged. A less complete version has appeared in CLS 14. Of course, all errors remain my own.

[1]In "Rappoteur's introduction and summary for the section on "Are Australian Languages Syntactically Nominative-Ergative or Nominative-Accusatice?" in Dixon (1976), Blake has the following remarks about the distri-

bution of ergativity in Australian languages: "The evidence available is not overwhelmingly strong, but the indications are that the typical Australian language is syntactically nominative-accusative. On the other hand Dyirbal is syntactically nominative-ergative, and it is the only Australian language that has been clearly documented as being ergative in syntax." (Dixon, 1976:488)

[2]For a more extensive listing of subject properties, see Keenan (1976).

[3]Johnson (1974) proposes the following Generalized Accessibility Hierarchy: Primary $\geq$ Secondary $\geq$ IO $\geq$ Oblique NP $\geq$ ... $\geq$ $\hat{\text{chomer}}$

[4]Dixon (1972) distinguishes an ergative case and an instrumental case although they are morphologically identical. Rather than maintain this distinction, the two kinds of NPs will be treated as being marked with the same case out as representing two grammatical relations. In Dyirbal instrumental (true instruments) and commitative NPs can be promoted to surface subject in the -m(b)al construction, but agent NPs in instrumental case cannot. This distinction is similar to the one between the ability of instruments, locatives, and other obliques, but not passive agents, to undergo passive under certain consitions. The fact that instruments but not agents can undergo -m(b)al promotion can be accounted for by both analyses. In an absolutive-ergative analysis the inability of the ergative to be promoted indicates that it is a distinct grammatical relation. However, since this implies that ergative NPs are very low on the Generalized Accessability Hierarchy, it is perhaps more problematical for the absolutive-ergative analysis than for the nominative-accusative analysis in which these non-promotable NP are treated as passive agents.

[5]The distinction between sentences like (6b) and (6c) is difficult to make explicit. In order not to misrepresent the situation, the following comment by Dixon (1972:66) is offered: "The difference between (examples parallel to (6c)) and (examples parallel to (6b)) is felt by speakers of Dyirbal to be a crutial and important one; it cannot be brought out through English glosses, but is explained in terms of the 'deep syntax'... Roughly, in (examples parallel to (6c)) the actor, goal, and action make up an event; (examples parallel to (6b)) imply something more -- that the actor is positively implicating the goal in the event. The difference is essentially one of topic."

[6]Use of the term Antipassive to characterize this kind of rule is perhaps misleading, especially for languages in which ergativity is a morphological phenomenon. In terms of grammatical relations, the rule is best referred to as Object Demotion. The effect of this demotion is to intransitivize the verb, which, in a morphologically ergative language, like West Circassian, results in absolutive (nominative) case marking on the subject and oblique (often identical with ergative) case marking on the demoted object. Note that when the transitive subject is marked with the same case as obliques, the surface morphology appears to have switched the cases of the absolutive NP, the direct object, and the ergative NP, transitive subject. Actually the subject is unaffected by the rule; the only change is the demotion of the object.

West Circassian (Anderson, 1976)

(i)p:sasa-m c'əy-ar ya-d-ə
girl-erg cherkessa-abs 3sg-(3sg)-sew-pres/trans
'The girl is sewing the cherkessa.'

(ii) p:sasa-r c'əy-əm ya-d-a
girl-abs cherkessa- obl 3sg-(3sg)-sew-pres/intrans
'The girl is sewing on the cherkessa.'

[7]This is not the only possible relationally based absolutive-ergative analysis. An alternative would be to posit two Antipassives, one demoting absolutives to obliques and another demoting absolutives to the second grammatical relation, ergative. What all relationally based treatments would have in common is that the ergative NP would represent the second highest relation and the dative case would not. Johnson (1974) tentatively proposes such a hierarchy of grammatical relations for an absoergative analysis. (See note 3.)

[8]George (1975) proposed an alternative analysis of Dyirbal in which he posits structures underlying sentences like (6b) as closest to the underlying representation of transitive sentences. This is all that his analysis has in common with the one which follows.

[9]Because the data for idiom arguments which could be used to establish which structure is underlying and which is derived is not available for Dyirbal, distinguishing between underlying and derived subjects will have to rely on other tests, such as Imperative Deletion. (See discussion of (19).)

[10]Imperative Deletion involves the pronoun system of Dyirbal, and the relationship of the surface case morphology to grammatical relations. (See discussion of (27).)

[11]In Dixon (1977) it is suggested that some transformations universally result in the same syntactic effect in every language in which they occur. Hypotheses regarding the possible kinds of 'split' syntactic ergativity fall essentially into two categories: those in which the syntax is split according to levels; either deep, shallow, surface, or cyclic, post and/or last cyclic, and those in which certain kinds of rules are sensitive to one of the possible sets of grammatical relations. The first position was basically that apopted by George (1975) in an earlier relational grammar treatment of Dyirbal; cyclic rules applied to a nominative-accusative syntax, then post cyclic rules applied to an absolutive-ergative syntax. However, in this analysis rules like Equi-NP-Deletion are treated as post-cyclic. The second position can be found in Woodbury's (1977) treatment of Greenlandic Eskimo where he posits a neutral NP hierarchy and general principals for determining the possible ranges of absolutive-ergative and nominative-accusative rules. Generally, absolutive-ergative relations are posited for morphological rules, rather than relation changing rules. The few exceptions to this include rules like Causative Clause Union and Antipassive (Object Demotion, see note 6). Close examination of these rules shows that not only is a neutral NP hierarchy unnecessary, but the notion of absolutive-ergative grammatical relations is unnecessary as well, since the only evidence for such relations is morphological; case marking and verb agreement rules can apply to the output of relation changing rules, as has been shown by Chung (1976a) and Harris (1976). Note that in a language like Georgian where the system of case marking is determined by tense, aspect, and verb classes, the grammatical relations of subject, object, indirect object, etc. only make sense in view of their syntactic behaviorial properties, not morphological coding properties.

[12] Note that in Dyirbal passives the pronouns referring to the passive subject and the passive agent are both in nominative case, but the derived passive subject is suffixed with -na/-ña. This parallels the morphology of passives in Imbabura Quechua in which the passive subject is obligatorily marked as a topic with -ca. This distinguishes it from the passive agent, a nominative case chômer. Note that in Imbabura Quechua the verb agrees with the derived subject in number and person.

(i) ñuca-ca pai-lla maca-shca ka-ni
1sg-top 3sg-just hit-pass be-1sg 'I was hit by him.'

(ii) kan-guna ñuca-chi yach-chi-shca ka-ngui-chi
2-pl 1-pl know-caus-pass be-2-pl 'You are taught by us.'

Because passive agents are not marked by instrumental case, there are two morphological forms of passive in Dyirbal. As shown above, in the pronoun system it is parallel to Imbabura Quechua. In the noun system the passive agent's demotion is indicated by instrumental case and the subject is optionally suffixed with -ña/-ña.

(i) ŋayguna ŋinda balga-n
1 sg-na 2 sg hit-non-fut 'I was hit by you'

(ii) bayi yara baŋgun ŋugumbiru balga-n
cm n man n cm i woman i hit-non-fut 'the man was hit by the woman'

For a language to have two forms of passive is not unprecedented. In Bahasa Indonesian (see Chung, 1976b) the distribution of the two passives seems to be similar to that of the two in Dyirbal; in general the non-canonical passive is used with first and second underlying subjects. Appartently this tendency may reflect an earlier stage in which the tendency was a restriction on third person underlying subjects; the canonical passive was supposed to have been used in these cases.

REFERENCES

Anderson, S.R. 1977. On mechanisms by which languages become ergative. In Li, ed.: Mechanisms of syntactic change. Austin: U Texas Press.

Anderson, S.R. 1976. On the notion of subject in ergative languages. In Li, ed.: Subject and topic. NY: Academic Press.

Blake, B.J. 1976. Rapporteur's introduction and summary. In Dixon, ed.: Grammatical categories in Australian languages. Canberra: Australian Institute of Aborigine Studies.

Catford, J.C. 1976. Ergativity in Caucasian languages. NELS 6.

Chung, S. 1976a. Case marking and syntactic structure in Polynesian. Harvard University dissertation.

Chung, S. 1976b. The subject of two passives in Bahasa Indonesian. In Li, ed.: Subject and topic. NY: Academic Press.

Comrie, B. forthcoming. Ergativity.

Dixon, R.M.W. 1972. The Dyirbal language of North Queensland. Cambridge University Press.

Dixon, R.M.W. 1977. The syntactic development of Australian languages. In Li, ed.: Mechanisms of syntactic change. Austin: U Texas Press.

George, L. 1975. Ergativity and relational grammar. NELS V.

Harris, A. 1976. Grammatical relations in Modern Georgian. Harvard University dissertation.

Horne, E. 1961. Beginning Javanese. New Haven: Yale University Press.

Johnson, D. 1974. Towards a theory relationally based grammar. University of Illinois dissertation.

Keenan, E.L. 1976. Towards a universal definition of 'subject'. In Li, ed.: Subject and topic. NY: Academic Press.

MacDonald, R. 1976. Indonesian reference grammar. Washington, D.C. Georgetown Press.

Woodbury, A. 1977. Greenlandic Eskimo, ergativity, and relational grammar. SS8.

Studies in the Linguistic Sciences
Volume 8, Number 1, Fall 1978

ON ERGATIVITY IN SELECTED SOUTH ASIAN LANGUAGES*

Yamuna Kachru
Rajeshwari Pandharipande

In a recent paper, it has been suggested that morphological ergativity may be "a consequence of syntactic ergativity and/or various types of semantic markings" (Dixon 1977, section 3.0). In this study, we examine data from a number of South Asian languages to determine (a) if these languages are syntactically and/or morphologically ergative and (b) in case they are morphologically ergative, if it is a consequence of syntactic ergativity or if it is conditioned by the semantic nature of NP's, verbs or Aspect-tense. We conclude that in the languages we have looked at, there is no evidence for syntactic ergativity. They may, however, be said to be morphologically ergative. But their morphological ergativity is not conditioned by any semantic marking and seems to be simply a consequence of the historical development.

1.0 In a recent paper (Pandharipande and Kachru 1977), we discussed the phenomenon of ergativity in Hindi-Urdu in the framework of Relational Grammar (Postal and Perlmutter 1976, Johnson 1974). In this paper, we attempt to look at ergativity in a wider context, namely, in the context of following South Asian languages: Bengali, Hindi-Urdu, Kashmiri, Marathi and Nepali. These languages represent both the inner (Hindi-Urdu) and outer group (Bengali and Marathi) of New Indo-Aryan (Chatterjee 1926), an Indo-Aryan language outside India (Nepali) and a Dardic language (Kashmiri). These languages provide us with a spread which is interesting both from theoretical as well as typological points of view. In particular, we examine whether ergativity in these South Asian languages is determined by the semantic nature of Noun Phrases (NP's, hereafter), Verbs (V's, hereafter), or Aspects-Tenses.

This study follows the format used in our earlier paper (Pandharipande and Kachru 1977) and proceeds as follows. First, we discuss morphological ergativity in the languages under consideration in terms of case (or postposition) marking and verbal agreement (section 2.0). Next, we determine if the morphological ergativity in these languages is "a consequence of syntactic ergativity and/or various types of semantic marking" (Dixon 1977, section 3.0). The final section states our conclusions.

2.0 A language is said to be ergative if the intransitive subject (SI, hereafter) is treated in the same manner as the transitive object (DO, hereafter), and differently from the transitive subject (ST, hereafter). The clearest evidence of such different treatment is in the case (or postposition/preposition) marking: in ergative languages, the ST is marked differently from both SI and DO. In the following sub-section, the marking of SI, ST and DO in each of the

languages under consideration is discussed separately. Any effect such marking has on verbal agreement is also pointed out.

2.1 In Bengali, neither the SI nor the ST or DO is marked with special inflection or postposition. Only an animate definite object is followed by the accusative/dative postposition ke. Consider the following sentences:

1. ami baṛT jacchT
   I home go 1st p.
   I am going home.

2. tumi baṛT jaccho
   you home go 2nd p.
   You are going home.

3. še baṛT jacche
   he/she home go 3rd p.
   He/She is going home.

4. amra baṛT jacchT
   we home go 1st p.
   We are going home.

5. chelera baṛT jacche
   boys home go 3rd p.
   The boys are going home.

6. amT boT poṛiam
   I book read 1st p.
   I read a book.

7. še boT poṛlo
   he book read 2nd p.
   He read a book.

8. mã cheleke ḍaklen
   mother boy acc. called hon.
   Mother called the boy.

9. še tomake ḍakchilo
   he you acc. call was 2nd p.
   He was calling you.

Notice that in each of the above examples, the SI as well as ST is unmarked and only the DO in 8-9 is marked with the accusative postposition ke. The verb shows agreement in person (1st, 2nd, 3rd), and honorific (sentence 8). We need not discuss the case of Bengali any further, as Bengali does not show any overt marking for ergativity. We shall, however, come back to Bengali later in our discussion (section 5.0).

2.2 Unlike Bengali, in Hindi-Urdu, the transitive verb in the perfective requires the agent to be marked with a special agentive

postposition ne, e.g., compare the following non-perfective and perfective sentences with intransitive and transitive verbs:

10. lərka sota hɛ / soya
boy sleeps / slept masc. sg.
The boy sleeps/slept.

11. lərki sotT hɛ / soT
girl sleeps / slept fem. sg.
The girl sleeps/slept.

12. mɛ̃ sota hũ
I sleep 1st p. masc. sg.
I sleep.

13. {lərke / lərkT} {boy / girl} ag. book read
The boy/girl read the book.

14. {lərke / lərkT} {boy / girl} ag. newspaper read
The boy/girl read the newspaper.

15. mɛ̃ ne kitab pərhT
I ag. book read
I read the book.

16. mɛ̃ ne əxbar pərha
I ag. newspaper read
I read the newspaper.

Notice that neither the SI in 10-12, nor the DO in 13-16 are marked with any postposition, whereas the ST in 13-16 is followed by the post-position ne. Notice further that in 10-12, the verb agrees with the SI, whereas in 13-16, the verb agrees with the DO, i.e., with kitab 'book' (fem. sg.) in 13 and 15, and əxbar 'newspaper' (masc. sg.) in 14 and 16. On the basis of data such as the above, Hindi-Urdu has been described as an ergative language (Allen 1950). This account of NP-marking and verbal agreement in Hindi-Urdu, however, is incomplete. Since the details of NP-marking and verbal agreement in Hindi-Urdu have been discussed elsewhere (Pandharipande and Kachru 1977), we will simply summarize them here. The following facts about Hindi-Urdu have to be taken into account before characterizing Hindi-Urdu as syntactically ergative. First, some transitive verbs in the perfective do not allow the agent marked with ne, (e.g., lana 'bring'), whereas others allow for optional marking with ne, (e.g., səməjhna 'understand'). Second, some intransitive verbs allow the SI to be marked with ne, (e.g., nəhana 'bathe oneself', chĩkna 'sneeze', etc.). Third, the verb agreement is not sensitive to ne, it is sensitive to any postposition. As a consequence, the verb either agrees with any unmarked NP in the sentence, or, if all NP's are marked, shows neutral agreement. Consider the following examples of such neutral agreement:

17. {lərke / lərkĩ} ne kitabõ ko uṭha liya
{boy / girl} ag. books acc. picked took
The boy/girl picked up the books.

18. mujhko bəccõ ko khilana hɛ
me dat. children acc. to feed is
I have to feed the children.

19. usse bəccõ ko səmbhala nəhĩ jaega
him by children acc. care not go will
He will not be able to care for the children.

Note that since all NP's in 17-19 are followed by postposition, the verb is in the unmarked, i.e., masc. sg. form. Note also that 18 and 19 are not perfective, rather, they are modal and passive, respectively, (see Pandharipande 1977 for a discussion of Passive in Hindi-Urdu).

Since the ne-marker is not determined either by transitivity or by perfective aspect, it is worth considering if the semantic nature of NP or V conditions the occurrence of ne in Hindi-Urdu (as suggested in Dixon 1977). Two possibilities are worth considering in this respect: ne always marks the agent as opposed to the patient or ne always marks the volitional agent as opposed to the non-volitional agent. Unfortunately, neither one of the above is borne out by the data. Compare the following:

20. ram (ne) chĩka
Ram (ag.) sneezed
Ram sneezed.

21. ram ko chĩk aĩ
Ram dat. sneeze came
Ram sneezed.

22. mɛ̃ (ne) khãsa
I (ag.) coughed
I coughed.

23. mujhko khãsĩ aĩ
me dat. cough came
I coughed.

It appears as though the dative subject construction characterizes the patient subject (in 21 and 23) whereas the agentive SI optionally takes ne (in 20 and 22). But, consider further data:

24. bəcca gira
child fell
The child fell down (accidentally or deliberately).

25. *bəcce ne gira
child ag. fell

26. ram ne təsvīr dekhī
Ram ag. picture saw
Ram saw/looked at the picture.

27. *ram təsvīr dekha
Ram picture saw.

28. ram kitab laya
Ram book brought
Ram brought the book.

29. *ram ne kitab laī
Ram ag. book brought.
Ram brought the book.

30. bariš ne sara məza kirkira kər diya
rain ag. all enjoyment ruined
The rain ruined all fun.

24-30 illustrate the fact that ne does not consistently mark the (volitional) agent in Hindi-Urdu. In dative subject construction, the dative postposition ko does mark the experiencer (Fillmore 1968) consistently (Kachru, Kachru, Bhatia 1976).

2.3 Nepali is similar to Hindi-Urdu in that the ST is marked with the instrumental postposition le. It is different from Hindi-Urdu in that the marking is not restricted to perfective only. Also, the verb agreement is not affected by le, the verb still agrees with the ST:

31. ma hāsẽ
I laughed.

32. keṭo hāso
boy laughed
The boy laughed.

33. maile kitāp paṛhẽ
I instr. book read
I read the book.

34. keṭale kitāp paṛho
boy instr. book read
The boy read the book.

The postposition le occurs with the third person ST in the future, and with all subjects (i.e., both SI and ST, in all persons) in the tenses formed with the infinitive (as in 37-40 below):

35. usle bhat khala
he instr. rice eat will
He will eat rice.

36. *usle ghara jālā
he instr. home go will
He will go home.

37. usle jānũ parchā
he instr. go inf. should
He should go.

38. usle kitāp parhnũ parchā
he instr. book read inf. should
He should read the book.

39. maile kitāp parhnũ parchā
I instr. book read inf. should
I should read the book.

40. usle / uslai jānũ cha
he instr./he dat. go imp. is
He has to go.

Notice that in case of modal constructions (37-40) the verb agreement is neutral, the verb does not agree either with SI or ST. In sentence 40, SI or ST may take either the instrumental le or the dative lai. A generic ST of an imperfective (habitual present) obligatorily takes le, as in the following:

41. rughāle dukha dincha
cold instr. suffering gives
Cold makes one suffer.

42. bāghle bākhrā khānchan
tigers ag. goats eat
Tigers eat goats.

43. *bāghle jangalma bascha
tiger ag. forests in live
Tigers live in forests.

To complicate matters further, le may occur with the ST in the perfective of verbs such as hācchiũ garnu 'sneeze', khoknu 'cough', etc.:

44. usle hācchiũ garo
he ag. sneeze did
He sneezed.

45. maile khokẽ
I ag. coughed
I coughed.

To summarize, the distribution of le in Nepali is as follows:

46. a. Simple past and perfective tenses: all ST, some SI
b. Future: 3rd person ST only
c. Present imperfect (habitual): generic ST only
d. Vnũ par-: all SI and ST (i.e., in all persons)
e. Vnũ ch-: all SI and ST (i.e., in all persons)
[In e, le~lai alternate]

Again, it is difficult to detect any consistent semantic motivation for this distribution.

2.4 As far as NP-marking is concerned, in Marathi, only the third person ST is marked with the instrumental postposition ne (variant nĩ) in the perfective, e.g.:

47. mĩ kame kelĩ
I jobs did 3rd pl. neuter
I did the jobs.

48. amhĩ kame kelĩ
we jobs did 3rd pl. neuter
We did the jobs.

49. tũ kame kelĩ(s)
you jobs did 3rd pl. neuter (2nd p.)
You did the jobs.

50. tumhĩ kame kelĩ(t)
you pl. jobs did 3rd pl. neuter (pl.)
You did the jobs.

51. tyane / tine kame kelĩ
he instr./she instr. jobs did 3rd pl. neuter
He/She did the jobs.

52. tyannĩ kame kelĩ
they instr. jobs did 3rd pl. neuter
They did the jobs.

Contrast the above sentences in the perfective with the following in the imperfective:

53. ram ambe khato
Ram mango eats (masc. sg.)
Ram eats mangos.

54. sudha ambe khate
Sudha mangos eats (fem. sg.)
Sudha eats mangos.

Except for the optional second person ending in 49 and 50, the verb agreement is similar to Hindi-Urdu. The NP-marking is different in that in Marathi, only the third person ST is marked with ne (nĩ). This is, however, an incomplete picture. In Marathi, the third person SI and ST are marked with ne (nĩ) in the desiderative and modal 'should'-construction also, e.g.:

55. ram ne ata ghərĩ zave
Ram instr. now home may go
Ram may go home now.

56. tyane gaṇṬ mhaṇṭlṬ pahijet
he instr. songs sing should pl.
He should sing songs.

Notice that the verb agrees with the DO in 56 and is in neutral agreement in 55, cf:

57. mṬ ata zave
I now may go
I may go now.

58. tyane / tumhṬ kam kərave
he instr. / you job may do
He/You may do the job.

As in the perfective, in the desiderative (55 and 57) and modal (56 and 58), too, the second person subject marking suffix may occur with the verb, e.g.:

59. tū zavas / zavṬs
You may go.

60. tu ghərṬ ala puhijes
you sg. home come masc. should
You should come home.

Note that in the modal construction, all subjects may optionally be marked with the dative -la, in which case, in the third person, we may get either -ne or -la, e.g.:

61. {məla / tula / tyala} ghəri alə pahije
{I / you / he} dat. home come should
I/You/He should come home.

The NP-marking and verb agreement pattern in Marathi seems to be as follows:

62. a. SI unmarked : V agrees with SI
b. ST unmarked in imperfective : V agrees with ST
c. ST marked (Ø with 1st/2nd, ne (nī) with 3rd) in perfective : V agrees with unmarked DO
d. ST marked (Ø with 1st/2nd, ne (nī) with 3rd) in desiderative/ modal : V agrees with DO if transitive, otherwise neutral
e. ST marked, DO marked : Neutral agreement

An example of e is as follows:

63. tyane mulṬla pahilə
he instr. girl acc. saw neut.
He saw the girl.

Note that Marathi is similar to Nepali as far as NP-marking is concerned: in both languages, subjects of perfective as well as certain modal constructions are marked. We will have more to say about this in 4.0.

2.5 In Kashmiri, the ST in the perfective is marked with specific case forms as opposed to the following which are in the nominative: the ST in the non-perfective, the SI in both perfective and non-perfective, and the DO in both perfective and non-perfective (Kachru 1969). Consider the following:

64. bɨ chus kh'ava:n
I be masc. sg. eat
I am eating.

65. tim chi kh'ava:n
they be masc. pl. eat
They are eating.

66. me kh'av batɨ
I dat. ate masc. sg. food masc. sg.
I ate food.

67. timav kheyi tsoṭ
they ag. ate fem. sg. bread fem. sg.
They ate bread.

The case forms utilized for marking the ST in the perfective are as follows (Kachru 1969):

68. First and second person : dative case
Third person: agentive case

The following sentences exemplify this:

69. rāman dits me kitāb
Ram ag. gave fem. sg. me book fem. sg.
Ram gave me a book.

70. me pər kitāb
I dat. read fem. sg. book fem. sg.
I read the book.

71. tse dits tamis tsoṭ
you dat. gave fem. sg. him dat. bread fem. sg.
You gave him bread.

72. təm' pər kitāb
he/she ag. read fem. sg. book fem. sg.
He/She read the book.

The dative (with first and second person) and agentive (with third person) is used with the pazi-imperative, too (Kachru 1969:268):

73. asi pazi kh'on
we dat. ought eat
We ought to eat.

Kashmiri has several nominal declension classes and very few nouns/pronouns have distinct forms for each case, hence, it is not unreasonable to label the so-called dative as dative/agentive in the context of first and second person pronouns. Note that mardav may be either the ablative or the agentive form of marɨd 'men'. Now, we can say that in Kashmiri, the ST in the perfective is marked with the agentive case-marker. The verb agreement pattern is similar to Hindi in that the verb agrees with the SI and also ST in the non-perfective tenses, while it agrees with the DO in the perfective tenses. The agreement pattern, however, is not as consistent as in Hindi. This is because, in Kashmiri, certain transitive verbs govern objects in oblique cases and others do not. For instance, the verbs 'see' and 'call' take objects in nominative case and therefore, the verb agrees with the DO; the verb 'hit', on the other hand, takes an oblique object, therefore, the verb is in its neutral form, i.e., it does not agree either with ST or DO.[1] Consider the following:

74. a. me bulov su ləḍkɨ
I dat. called that boy nom.
I called that boy.

b. me bulə:v sɔ kūr
I dat. called that girl nom.
I called that girl.

75. me loy təmis ləḍkɨ
tas kori
I hit that boy dat.
that girl
I hit that boy/that girl.

In 74 a and b, the verb bulov versus bulə:v shows agreement with the DO; in 75 the verb form is in third person masculine singular form.

3.0 The above discussion suggests that except for Bengali, all the South Asian languages under study may be at least morphologically 'split ergative' (Dixon 1977). In this section, we will determine if any of the above languages are syntactically ergative, too.

According to Dixon 1977, there are three universal syntactico-semantic primitives - A and O occurring in transitive sentences and S occurring in intransitive sentences. 'Subject' is a universal deep structure category, involving functions A and S. The operation of optional singulary transformations (passive, antipassive, reflexive, etc.) on deep structures yields shallow structures. It is at this level that generalized transformations operate, forming coordinate and subordinate constructions. These rules may treat (derived) S and A in the same way (i.e., S/A pivot) or they may treat (derived) S and O in the same way (i.e., S/O pivot). If a language utilizes the S/O pivot, it can be said to have an 'ergative' syntax.

3.1 We have already shown elsewhere that neither Hindi-Urdu, nor Kashmiri work in terms of a S/O pivot (Kachru, Kachru and Bhatia 1976), hence there is no evidence for syntactic ergativity in these languages.

Sharma 1977 presents evidence to show that Nepali uses a S/A pivot in major subordinate and coordinate structures. Pandharipande 1978 presents similar evidence for Marathi. The following is a summary of the evidence presented in the works mentioned above.

3.2 First, in terms of the accessibility hierarchy (Keenan, Comrie 1972), the marked ST is as accessible to relativization as SI or DO or any oblique NP in the South Asian languages under study here. The following exemplify this (H-U, K, M, N preceding the numbers stand respectively for Hindi-Urdu, Kashmiri, Marathi and Nepali):

H-U 76. jis lərke ne kitab xərīdī thī vəh rezgarī lena bhūl gəya
wh- boy ag. book bought had he change take forget went
The boy who bought the book forgot to take the change.

K 77. su ləḍkɨ yem rāth bə̄th gʲov čhu yeti rōzān
that boy who yesterday song sang is here living
The boy who sang the song lives here.

M 78. jya mulane kal gaṇe mhəṭle to ithec rahto
which boy yesterday song sang he here lives
The boy who sang the song yesterday lives here.

N 79. jun keṭāle hijo git gāyo tyo yāhī bascha
which boy yesterday song sang that (boy) here lives.
The boy who sang the song yesterday lives here.

3.3 The conjunction reduction rule which reduces the first conjunct of a coordinate structure by deleting the identical subject and transforming the finite verb into an absolutive form, utilizes S/A pivot in the languages under consideration here.[2] Note, however, that whereas in Hindi-Urdu, Kashmiri and Marathi, DO does not either control or undergo deletion; in Nepali, DO undergoes deletion even though it does not control deletion. For instance:

H-U 80. mɛ̃ ne khana khaya ɔr mɛ̃ so gəya
I ag. food ate and I sleep went
I ate (some) food and went to sleep.

H-U 81. khana kha kər mɛ̃ so gəya
food eat having I sleep went
Having eaten, I went to sleep.

H-U 82. mɛ̃ ne bəcce ko mara ɔr bəcca roya
I ag. child acc. hit and child cried
I hit the child and the child cried.

H-U 83. *(mɛ̃) mar kər bəcca roya
I hit having child cried
I having hit (him), the child cried.

K 84. me kʰ'av batɨ tə šɔ̃ngus
I food ate and slept
I ate and went to sleep.

K 85. bɨ šɔ̃gus batɨ kh'ath
I slept food having eaten
or
batɨ kh'ath šɔ̃gus bɨ
food having eaten slept I
Having eaten, I went to sleep.

K 86. me lo:v ləḍkas tɨ təm' h'ot radun
I hit child and he cried
I hit the child and he cried.

K 87. *me mə:rith h'ot ləḍkan radun
I having hit child cried
I having hit, the child cried.

M 88. mɪ̃ jevlo aɳɪ̃ mɪ̃ zoplo
I ate and I slept
I ate and I slept.

M 89. Ø jevũn mɪ̃ zoplo
having eaten I slept
Having eaten, I went to sleep.

M 90. mɪ̃ mulala marlə aɳɪ̃ mulga rəḍla
I boy obj. masc. hit and boy cried.
I hit the boy and the boy cried.

M 91. *mɪ̃ mulala marũn mulga rəḍla
I boy obj. masc. having hit boy cried.
I having hit (him) the boy cried.

N 92. maile gulivorko yātrā paṛhẽ ani ma hāsẽ
I Gulliver's travels read and I laughed
I read Gulliver's Travels and I laughed.

N 93. gulivarko yātrā paṛhera ma hāsẽ
Gulliver's travels having read I laughed
Having read Gulliver's Travels, I laughed.

N 94. usle chorilāi kuṭo ra chori roi
he instr. girl acc. hit and girl cried
He hit the girl and the girl wept.

N 95. usle kuṭera chori roi
he instr. hit having girl wept
He having hit (her), the girl wept.

N 96. chori roi ra usle chorilāi kuṭo
girl wept and he instr. girl acc. hit
The girl wept and he hit the girl.

N 97. a. *roera usle chorilāi kuṭo
having cried he instr. girl acc. hit
(The girl) having cried, he hit the girl.

b. *chori roera usle kuṭo
girl having cried he instr. hit
The girl having cried he hit (her).

3.4 The rule that yields past participle modifiers of NP's in Hindi-Urdu seems to be sensitive to ergative/absolutive distinction, for example:

H-U 98. [lərkiyã bəṭhT̃] lərkiyã
girls sat girls

H-U 98a. bəṭhT huT lərkiyã
The seated girls.

H-U 99. [lərkiyõ ne kitab pərhT] lərkiyã
girls ag. book read girls

H-U 99a. *kitab pərhT huT lərkiyã
The book-read girls.

H-U 100. [lərkiyõ ne kitab pərhT] kitab
girls ag. book read book

H-U 100a. lərkiyõ kT pərhT huT kitab
The read-by-the-girls book.

Notice that 98a and 100a are grammatical whereas 99a is ungrammatical: in 98a, the past participial modifies an NP identical to the S-NP, in 100a, it modifies an NP identical to DO of the participial verb; in 99a, were it grammatical, the participial would modify an NP identical to ST. It has, however, been shown (Pandharipande and Kachru 1977) that the facts about participialization in Hindi-Urdu are to be explained on the basis of the semantic nature of NP's and verbs rather than ergative/absolutive distinction.

In Nepali, the rules that yield past participial and infinitival participial seems to utilize the S/O pivot:

N 101. [keṭiharule kam gare] kam
girls instr. work did work

N 101a. keṭiharule gareko kam
The done-by-the-girls work

N 102. [keṭiharule kam gare] keṭi
girls instr. work did girls

N 102a. *kam garekā keṭiharule
The work-done girls.

N 103. [mānche maro] mānche
man died man

N 103a. mareko mānche
The dead man.

N 104. [keṭale kitāp paṛh] kitāp
boy instr. book read book

N 104a. keṭale paṛhne kitāp
boy instr. to read book
The book to be read by the boy.

N 105. [keṭale kitāp paṛh] keṭa
boy instr. book read boy

N 105a. *kitāp paṛhne keṭa
book to read boy

N 106. [mānche hāse] mānche
man laugh man

N 106a. hāsne mānche
The man to laugh

Not much information is available on participialization in Marathi and Bengali. As far as Bengali is concerned, participialization seems to be as marginal a process as it is in English (Verma 1968). Marathi seems to be more like the Dravidian languages where a wide range of participial modifiers are used. There is no evidence at present to suggest that either one of these languages utilizes a S/O pivot for deriving participial modifiers.

3.5 The discussion so far makes it clear that except for Nepali, which uses S/O pivot for **deriving participial modifiers, there is no** evidence for syntactic ergativity in any other South Asian language under discussion here. There is no rule such as antipassive in any of these languages, and the passive in South Asian languages does not necessarily yield a derived subject (Kachru, Kachru and Bhatia 1976). As far as participles are considered, the case of Hindi-Urdu (Pandharipande and Kachru 1977) suggests that a deeper examination of language-internal facts is necessary before arriving at any conclusions. What appears to depend upon S/O pivot in Nepali may turn out to have an alternate explanation in Nepali, too.

4.0 In conclusion, it is necessary to evaluate all the evidence we have presented so far. In the absence of any evidence for syntactic ergativity, the question naturally arises as to whether South Asian languages exhibit morphological ergativity. On the surface, the languages under discussion here, except Bengali, seem to be morphologically ergative. Hindi-Urdu has a special postposition ne to mark A as opposed to S and O, Kashmiri has a special agentive case, and Marathi and Nepali utilize the instrumental postposition to mark the A. The marking of third person A only in Marathi and partially in Nepali also follow the principles suggested by Dixon 1977. That is, on the 'potentiality of agency scale', third person forms rank lower as compared to first and second person forms. The marking of A in perfective also follows the principles suggested in Dixon 1977. Note, however, that Marathi and Nepali (and marginally, Kashmiri) go beyond the past - non-past distinction suggested in Dixon 1977. The principle utilized by these languages suggests the following type of split:

107a. Nepali: Now versus non-now (timeless and potential)
b. Marathi: Past versus non-past, and actual versus potential

Kashmiri is closer to Marathi in this respect. Even this, however, does not explain the marking exhaustively. If 107a and b were to be consistently applicable, imperative, and subjunctive forms should have fallen together with the modal constructions in these languages. That does not seem to have happened.

In the absence of any synchronic explanation, it seems best to rely upon a diachronic explanation. The perfective in New Indo-Aryan is said to have developed from the past passive participle of old Indic (Chatterjee 1926). That explains the instrumental marking of the A in Marathi and Nepali. The agentive postposition and case in Hindi-Urdu and Kashmiri respectively seem to be parallel developments, too. It, however, raises a very interesting question with regard to Bengali. Notice that Bengali went through identical historical stages. Nevertheless, instead of an NP-marking system (for A), what developed in Bengali was a system of verbal affixation which copied on to the verb the personal pronouns. Thus, both S and A are unmarked and control the verbal agreement. Several dialects of the so-called Hindi-area went through identical development and the tendency in Eastern as well as Dakhini varieties is toward the same kind of development. Western Hindi, however, has the stable ne-marker, and Marathi seems to be stabilizing the O-V agreement in the transitive sentences by losing the optional second person suffix of the verb in sentences such as 49, repeated here for convenience:

49. tū kame kelī (s)
you jobs did 3rd pl. neut. (2nd p.)
You did the jobs.

The verbal agreement pattern does not show a neat split, however, in that all the languages under consideration mark the O in specific linguistic contexts.

Dixon 1977 is no doubt a valiant attempt to bring some system to the chaos. It, however, seems doubtful, on the basis of data from South Asian languages, that it is always possible to explain morphological ergativity on the basis of semantic features of NP, V, or Aspect-Tense. The languages we have discussed here, except for Bengali, seem to be part accusative, part ergative and part neither and these parts seem to be determined by historical development rather than consistent semantic principles. This suggests that the inventory of types of split in Dixon 1977 should include a fourth category: split conditioned by historical syntactic/morphological development.

NOTES

*Part of the research reported herein was carried out with the help of a research grant from the Research Board of the University of Illinois. The work on Marathi was done primarily by Rajeshwari Pandharipande and the Nepali data was supplied by Basudev Sharma.

[1]In Kashmiri, as in Hindi-Urdu, the verb agrees with the noun in the direct or nominative case. In the perfective, the ST is in the

oblique case, (i.e., a case other than the nominative) and the DO is in the nominative, therfore, the verb agrees with the DO. In case a particular verb governs a DO in an oblique case, in the perfective, both the ST and the DO are in non-nominative cases, hence the verb agreement is neutral. This has been discussed in detail in Kachru, Kachru, and Bhatia 1976.

[2]The conjunction reduction rule discussed here works as follows. Given (i) and (ii), the rule yields (iii):

H-U (i) ram ghər gəya
Ram home went
Ram went home.

H-U (ii) ram ne khana khaya
Ram ag. meal ate
Ram ate (his) meal.

H-U (iii) ghər ja kər ram ne khana khaya
home go abs. Ram ag. meal ate
Having gone home, Ram ate (his) meal.

In Hindi-Urdu, there are only two constructions with absolutive forms which do not require the identity and coreferentiality of the S/A in the two clauses. These are as in (iv) and (v):

H-U (iv) car baj kər dəs minət hue hɛ̃
four strike abs. ten minutes happened are
It is ten past four.

H-U (v) bhaT ko dekh kər uska mən šant hua
brother DO see abs. his mind calm became
He was reassured to see his brother.

Sentences (iv) and (v) involve the sets given in (vi) and (vii) respectively:

H-U (vi) a. car bəje hɛ̃
four struck is
It is four.

b. dəs minət hue hɛ̃
ten minutes happened are
Ten minutes have passed.

H-U (vii) a. usne bhaT ko dekha
he ag. brother D saw
He saw (his) brother.

b. uska mən šant hua
his mind calm became
He was reassured.

Note that apparently, in (vi), the two S's are distinct, and in (vii) the two A's are not identical: in (viia), the A is He, in (viib),

it is His mind. The A in (viia), however, is coreferential with the possessive NP of A in (viib). It is interesting to note that in Hindi-Urdu, other processes sensitive to coreferential relations behave similar to the conjunction reduction as in (viib).

REFERENCES

Allen, W.S. 1950. A study in the analysis of Hindi sentence structure. Acta Linguistica Hafniensin 6. pp.68-86.

Chatterjee, S.K. 1926. The origin and development of the Bengali language. London: Allen and Unwin.

Comrie, B. 1976. (rev. 1977). Ergativity. (mimeographed).

Dixon, R.M.W. 1977. Ergativity. (mimeographed).

Fillmore, C.J. 1968. The case for case. In E. Bach and R.T. Harms, eds.: Universals in linguistic theory, pp. 1-90. Holt.

Kachru, B.B. 1969. A reference grammar of Kashmiri. Department of Linguistics, University of Illinois.

Kachru, Y., B. Kachru and Tej Bhatia. 1976. The notion 'subject': a note on Hindi-Urdu, Kashmiri and Punjabi. In M.K. Verma, ed.: The notion 'subject' in South Asian languages, pp. 79-108. Madison, University of Wisconsin.

Keenan, E. and B. Comrie. 1972. Noun phrase accessibility and universal grammar. Linguistic Inquiry 8:1, 1977. 63-100.

Pandharipande, R. and Y. Kachru. 1977. Relational grammar, ergativity and Hindi-Urdu. Lingua 41. pp. 217-38.

Pandharipande, Rajeshwari. 1977. Exceptions and rule government; the case of the passive rule in Hindi. Paper presented at the 187th meeting of the American Oriental Society held at Ithaca.

_____. 1978. Participles in Marathi. (unpublished).

Sharma, Basudev. 1977. The syntax and semantics of participles in Nepali. (unpublished).

Verma, M.K. 1971. The structure of the noun phrase in English and Hindi. Delhi. Motilal Banarasidas.

Studies in the Linguistic Sciences
Volume 8, Number 1, Spring 1978

THE RELATIVE PRONOUN AND THE LONG FORM ADJECTIVE IN RUSSIAN

James Levine
University of Maryland

Babby's (1973, 1975) analysis of Russian attributive adjectives is critically discussed. A reanalysis is proposed on the basis of the assumption that J and not KOTOR is the relative pronoun in Russian.

The syntactic derivation of adjectives is probably one of the less disputed issues within the theory of generative grammar. There is wide agreement in the assumption that attributive adjectives are related to underlying predicates (Harris 1952; Chomsky 1957; Lees 1960a), and that they are transformationally derived from deeper relative clauses, i.e., from embedded sentences (Smith 1961, 1964; Lees 1960b; Katz and Postal 1964; Lakoff 1970; Ljung 1970, etc.). While there are difficulties in the analysis deriving attributive adjectives from predicate adjectives in relative clauses (Bolinger 1967; Motsch 1967; Sussex 1974), a more promising explanation has yet to be found for the obvious relationship between adjectives in attributive and predicative positions.

The relative clause source of attributive adjectives is given additional support in recent studies of Russian adjectives by Babby (1973, 1975). At the center of Babby's investigations is the syntactic relation between the attributive long form (LF) and predicative short form (SF) of adjectives in Contemporary Russian.[1] In particular, he argues that the LF and SF are surface structure categories which derive from a single lexical category, Verbal (V). Under this analysis, what is traditionally called "adjective" is V with the feature [+adj], and what is traditionally called "verb" is V with the feature [-adj] (cf. Lakoff 1970 : Appendix A). Neither the SF nor the LF are assumed to be generated by the phrase structure rules. Rather, it is argued that the deep structure category V emerges from the transformational component as a SF if it acquires the features of gender, number and person by the Subject-Verb Agreement transformation. The category V emerges from the transformational component as a LF, if in addition to acquiring these agreement features, it receives a case feature. Babby states his main hypothesis as follows (1975:9): "the LF is a SF that has acquired a case feature by virtue of its transformational introduction into the constituency of a NP (noun phrase)." This analysis claims that it is the derived phrase marker configuration $[_{...}V_{...}]_{NP}$ which is the defining criterion of the LF. Furthermore, it assumes, in accordance with the "lexicalist" theory, that inflectional morphology is

handled by a set of postsyntactic morphophonemic rules which "spell out" features acquired in derivation as inflectional endings. For example, within this framework, the ending -AJA of the LF KRASIV-AJA in the NP KRASIVAJA DEVUŠKA "the pretty girl" is assumed to be a "spelling" of the features assigned by rules of agreement (fem, sing, 3rd) and case marking (nom). It is the opinion of the present writer that assuming the operation of such vague morphophonemic spelling rules does not adequately account for the morphological details of inflection in the LF adjective. The aim of this paper, therefore, is to present an alternative analysis in which some concrete, albeit tentative proposals are made as to how the syntax might account in a more explicit fashion for the inflectional endings of the LF.

Descriptions of Contemporary Russian by linguists operating within both traditional and generative frameworks assume the relative pronoun to be the morpheme KOTOR-, plus inflectional endings. For example, in the generative account of Russian adjectives in Babby 1975, the structure

$$\underset{\text{'girl'.}}{[[\text{DEVUŠK-}]\ \text{NP}}.\ \underset{\text{'girl'}}{[[\text{DEVUŠK-}]\ \text{NP}}\ \underset{\text{'is pretty'}}{[\text{KRASIV-}]\ \text{VP}]\ \text{S}]\ \text{NP}}$$

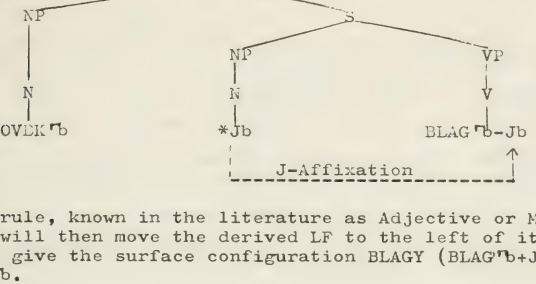
which is posited as the underlying representation of the NP KRASIVAJA DEVUŠKA 'pretty girl' undergoes Relative-clause Formation to produce the intermediate configuration

$$[\underset{\text{'girl'}}{[\text{DEVUŠK-}]\ \text{NP}}\ \underset{\text{'who'}}{[\text{KOTOR-}]\ \text{NP}}\ \underset{\text{'is pretty'}}{[\text{KRASIV-}]\ \text{VP}]\ \text{S}]\ \text{NP}}$$

In the present analysis I would like to propose that what is traditionally considered a morphologically whole inflectional ending in LF adjectives, actually contains the relative pronoun morpheme J, and that KOTORYJ is a surface configuration, transformationally derived by the prefixation of KOTOR- to this same morpheme J.[2] In arguing that the relative pronoun is J, I am in effect suggesting that one of the rules underlying the formation of the LF adjective in Contemporary Russian is the same rule assumed to have been operative in Old Russian. As is well known, in Old Russian the relative was the morpheme J (J-b nom masc) followed by the suffixed particle -ŽE. The traditional analysis of the formation of the LF adjective in Old Russian is that of Vaillant 1942, who argues that the relative pronoun eventually became affixed to a regular (i.e. short) adjective to create the LF. In assuming the concatenation of short adjectives and J, Vaillant provides a plausible explanation for the synonymous readings of Adj+N phrases and corresponding relative clauses. Thus, in Vaillant's view, the NP BLAGY (=BLAG⁀b+Jь) ČLOVEK⁀b 'the good man' and the relative clause ČLOVEK⁀b IŽE (=*Jь-ŽE) BLAG⁀b 'the man who is good' are identical constructions (p.5). If one were to incorporate this insight in a generative-transformational treatment of the LF in Old Russian, then Adj(LF)+N phrases would be analyzed as deriving from the deep structure

underlying corresponding relative clauses. Under such an analysis, a transformational rule which I shall call J-Affixation, would detach the relative pronoun morpheme from the N-node and affix it to the right of the SF, as in (1).

(1)



A late rule, known in the literature as Adjective or Modifier Shift, will then move the derived LF to the left of its head-noun to give the surface configuration BLAGY (BLAGˈb+Jb) ČLOVEKˈb.

The view that J is still the relative pronoun morpheme in Contemporary Russian and that the proposed rule of J-Affixation is still operative, finds support from the following morphological and syntactic considerations. Zaliznjak (1967:2329) argues that the inflectional endings of attributive adjectives in Russian contain a morpheme which functions as the marker of long formedness. His analysis is based on an examination of LF adjectives and their corresponding SF's:

ZIV-OJ	ŽIV-AJA	ŽIV-OJO	ŽIV-IJI
ŽIV	ŽIV-A	ŽIV-O	ŽIV-I

Comparing the long and short forms above, Zaliznjak notes that the LF adjective ending consists of a vowel identical to the vowel in the corresponding SF ending plus a 'special morphological element' J, followed by a vowel which is a copy of the one immediately preceding this element.[3] He calls the morpheme J the 'marker of attributive function'. If one accepts Zaliznjak's view that the inflectional endings of LF adjectives in Contemporary Russian are complex items of the structure V^1+J+V^1, where J is the long formedness morpheme, then, in my view, it is plausible to assume that this morpheme J is the same relative pronoun J which Vaillant claims was affixed to the SF to create the LF in Old Russian.[4]

Adopting the position that the relative pronoun in Contemporary Russian is J makes it possible to explain certain syntactic facts about attributive adjectives and corresponding relative clauses

with predicative adjectives. Consider, for example, sentences (a)-(c) below, taken from Babby 1973:

(a) MAL'ČIK, KOTORYJ BOLEN (SF) ANGINOJ, DOLŽEN LEŽAT'
'The boy who is ill with quinsy has to stay in bed
CELUJU NEDELJU.
all week.'

(b) MAL'ČIK BOL'NOJ (LF) ANGINOJ, DOLŽEN, LEŽAT' CELUJU NEDELJU
'The boy ill with quinsy has to stay in bed all week.'

(c) *MAL'ČIK, KOTORYJ BOL'NOJ ANGINOJ, DOLŽEN LEŽAT'...
'The boy, who ill with quinsy has to stay in bed ...'

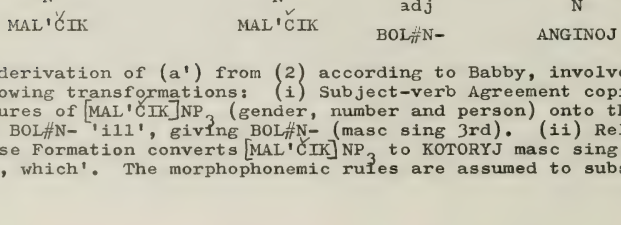
Sentence (a) consists of a relative clause introduced by KOTORYJ 'who', and contains the predicate adjective (SF) BOLEN 'ill'. Sentence (b) is a reduced version of sentence (a), containing BOL'NOJ 'ill' (LF) but not the relative pronoun. Sentence (c), marked deviant, contains both KOTORYJ and BOL'NOJ (LJ). Babby proposes an analysis of sentences (a) and (b), but does not address the question of why sentence (c) is ungrammatical. I shall repeat his analysis here and then propose an alternative analysis which will not only explicate the relationship between sentences (a) and (b), but will also account for the deviancy of sentence (c).

The deep structure proposed by Babby for the NP's in (a') and (b') is (2).

(a') [MAL'ČIK, KOTORYJ BOLEN ANGINOJ] NP
'the boy who is ill with quinsy'

(b') [MAL'ČIK, BOL'NOJ ANGINOJ] NP
'the boy ill with quinsy'

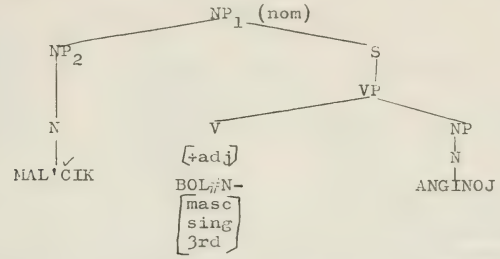
(2)



The derivation of (a') from (2) according to Babby, involves the following transformations: (i) Subject-verb Agreement copies the features of [MAL'ČIK]NP_2 (gender, number and person) onto the main verb BOL#N- 'ill', giving BOL#N- (masc sing 3rd). (ii) Relative-clause Formation converts [MAL'ČIK]NP_3 to KOTORYJ masc sing nom 'who, which'. The morphophonemic rules are assumed to subsequently

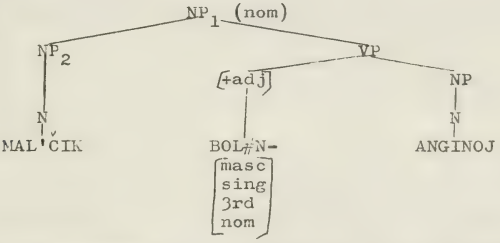
spell out BOL#N- (masc sing 3rd) as BOLEN (SF). The derivation of (b') from (2) involves both rules (i) and (ii) plus the optional transformation of Relative-clause Reduction, which results in the intermediate phrase marker of (3).

(3)



(3) is mapped into (4), the final derived phrase marker, by the rule of S-node Deletion which deletes the node S and results in the introduction of BOL#N- (masc sing 3rd) into the constituent of NP_1 where it receives its case feature (nominative) by the rule of Case Marking. Again it is assumed that the morphophonemic rules will later spell out BOL#N- plus the features acquired by Subject-verb Agreement and Case Marking, as BOL'NOJ (LF) masc sing nom 'ill'.

(4)

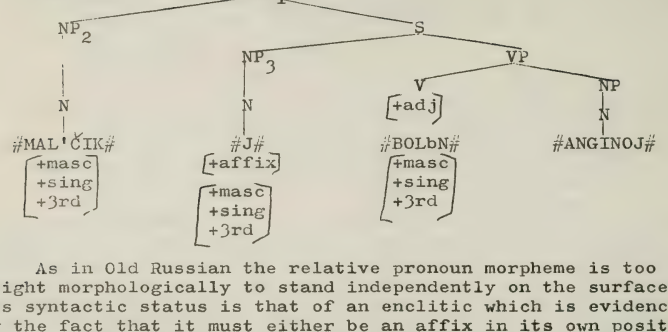


Babby's transformational treatment of the SF and LF adjective has served both as a model and catalyst for the analysis I am about to propose. In fact the two analyses differ essentially in one small, but very significant aspect, viz., the assumption that J and not KOTOR- is the relative pronoun morpheme in Contemporary Russian.

The proposed base structure for the NP's in (a') and (b') in the present analysis is (2), the structure proposed by Babby. As in Babby's analysis, the derivation of (a') involves both Subject-verb Agreement and Relative-clause Formation, except that in the present analysis, the latter rule relativizes

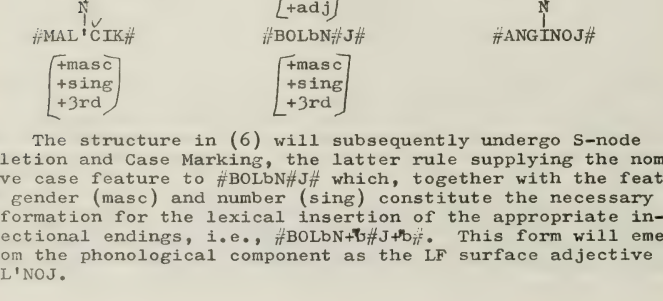
[MAL'ČIK]NP_3 to J, giving (5).[5]

(5)

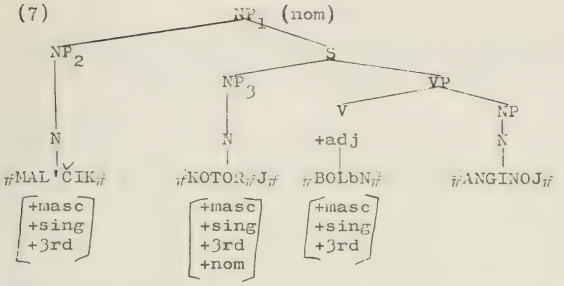


As in Old Russian the relative pronoun morpheme is too slight morphologically to stand independently on the surface. Its syntactic status is that of an enclitic which is evidenced by the fact that it must either be an affix in its own position or be affixed to some other lexical item. Thus, (5) represents the stage in the derivation at which one of two rules may be invoked: either the rule of J-Affixation that attaches J to the lexical item at the V-node (#BOLbN#) to allow for the creation of the LF (#BOLbN#J#), or the rule of KOTOR-Support is obligatorily invoked, inserting the formative KOTOR- to the left of J, giving #KOTOR#J#. The phrase markers resulting from the application of the rules of J-Affixation and KOTOR-Support are represented in (6) and (7) respectively.

(6)



The structure in (6) will subsequently undergo S-node Deletion and Case Marking, the latter rule supplying the nominative case feature to #BOLbN#J# which, together with the features of gender (masc) and number (sing) constitute the necessary information for the lexical insertion of the appropriate inflectional endings, i.e., #BOLbN+ъ#J+ъ#. This form will emerge from the phonological component as the LF surface adjective BOL'NOJ.



The structure in (7) undergoes no additional syntactic transformations. The case feature needed for the selection of the appropriate inflectional endings by the NP #KOTOR#J#, is supplied by the subject position of this NP in the embedded clause. Thus, the features nom, masc, sing will provide the specification needed for the lexical insertion of inflectional endings, giving #KOTOR+b#J+b#. The rules of phonology will convert this form to the surface relative pronoun, KOTORYJ.

A comparison of this analysis with Babby's suggests that both analyses account equally well for the formation of the LF adjective. However, in my view, the alternative analysis presented in this paper has the further advantage of offering a more explicit explanation for why sentences like (c) *MAL'ČIK, KOTORYJ BOL'NOJ ANGINOJ ... containing both the surface relative pronoun and a LF adjective are ungrammatical. Simply stated, if J is the same morpheme that occurs in both KOTORY(J) and BOL'NO(J), then the two forms cannot co-occur in the same clause. In syntactic terms, the rule of J-Affixation is complementary with the rule of KOTOR-Support.

It has been argued above that the LF adjective is not merely a SF that has acquired the features of gender, number and case which later get spelled out as inflectional endings. Rather, it has been shown that a LF is a SF which, in addition to these agreement features, has undergone concatenation with the relative pronoun morpheme by the rule of J-Affixation. Therefore, what has traditionally been considered a morphologically whole inflectional ending has been shown in this paper to be a complex item consisting of both derivational and inflectional formatives.

NOTES

[1]Babby's treatment of participles in these studies will not be considered in this paper.

[2]For an earlier treatment of this question, see Levine (1977: Ch. 4).

[3]According to Zaliznjak, there is no anomaly in the masculine form, since in this instance the O preceding the J is merely a "vocalic infix" which breaks up the inadmissable consonant cluster "consonant + J." Phonemic notation is employed here.

[4]Zaliznjak's discussion is limited to adjectives in the nominative case. In fact all of the nominative and accusative forms as well as all the feminine forms in the oblique cases of LF adjectives show a J in their surfaces phonetic representation. However, this analysis derives less support from those oblique case endings where there is no evidence of a J, i.e., in the masculine, neuter and plural forms. It should be pointed out that this is the same morphophonemic complication which obtains in Old Russian and which has resisted satisfactory historical explanation. One possible way to account for the absence of J would be to assume that certain oblique case forms are subject to a transformational rule which deletes this morpheme. For a discussion of this problem and a proposed deletion rule of this type, see Coats' description of the inflectional morphology of LF adjectives in Old Russian (Coats 1973).

[5]I have altered Babby's representation of the vowel/zero alternation in the SF adjective, replacing # with b, so as not to confuse this use of # with its use as a boundary symbol.

A few words should be said about the notational conventions employed in this description. In accordance with Chomsky and Halle (1968:364-369), I shall assume that each formative in the lexicon is flanked by a formative boundary symbolized by the plus sign +, e.g., +X+, +Y+, +Z+. I shall further assume, following Chomsky and Halle, that the formative boundaries of lexical items inserted at the terminal nodes of phrase markers are deleted when adjacent to the word boundary symbol #, which, by convention, is automatically introduced at the beginning and end of every string dominated by a major lexical category, e.g., #+X+# → #X#. The concatenation of two formatives resulting from a rule of affixation will produce a string with an initial and final #, but with only one internal +. For example, the concatenation of #X# and #Y# will give #X+Y#. However, the concatenation of the relative pronoun enclitic J with a lexical item will be represented slightly differently. The syntactic status of J will be characterized by the presence of a single internal word boundary symbol #, between the lexical item and the enclitic J. Thus, when J is affixed to the complex item #X+Y#, the resulting new lexical item will be represented as #X+Y#J#. Finally, I will assume that inflectional endings are inserted to the left of the right word boundary of a formative, and in the case of a formative to which J has been affixed, to the left of both right word boundary symbols, e.g., #X+Y+E#J+E#, where E stands for inflectional ending.

REFERENCES

BABBY, Leonard H. 1973. The deep structure of adjectives and participles in Russian. Language 49: 349-360.

_____. 1975. A transformational grammar of Russian adjectives. The Hague: Mouton.

BOLINGER, Dwight L. 1967. Adjectives in English: attribution and predication. Lingua 18: 1-34.

CHAPIN, Paul G. 1967. On the syntax of word derivation in English. MIT Ph.D. Dissertation.

CHOMSKY, Noam. 1957. Syntactic structures. The Hague: Mouton.

_____ and Morris Halle. 1968. The sound pattern of English. New York: Harper and Row.

CHOMSKY, Noam. 1970. Remarks on nominalization. In Roderick A. Jacobs and Peter S. Rosenbaum (eds.), Readings in English transformational grammar. Waltham: Ginn.

COATS, Herbert S. 1973. Old Russian declension: a synchronic analysis. In American contributions to the seventh international congress of slavists. Vol. 1, linguistics and poetics. ed. Ladislav Matejka. pp. 67-99.

HALLE, Morris. 1973. The accentuation of Russian words. Language 49: 312-48.

HARRIS, Zellig S. 1952. Discourse analysis. Language 28: 18-23.

JACKENDOFF, Ray. 1975. Morphological and semantic regularities in the lexicon. Language 51: 639-71.

KATZ, Jerrold and Paul M. Postal. 1964. An integrated theory of linguistic descriptions. Cambridge, Mass.: MIT Press.

LAKOFF, George. 1970. Irregularity in syntax. New York: Holt, Rinehart & Winston.

LEES, Robert B. 1960a. The grammar of English nominalizations. Bloomington: Indiana University.

____. 1960b. A multiply ambiguous adjectival construction in English. Language 36: 207-221.

LEVINE, James S. 1977. Towards a transformational account of derived adjectives in Russian. University of Illinois Ph.D. Dissertation.

LJUNG, Magnus. 1970. English denominal adjectives. Gothenburg, Sweden: University of Göteborg (Gothenburg Studies in English, 21.)

MOTSCH, W. 1967. Können attributive Adjektive durch Transformationen erklärt werden? Folia Linguistica 1: 23-48.

POSTAL, Paul M. 1964. Limitations of phrase structure grammars. In Jerry Fodor and Jerrold Katz (eds.), The structure of language. Englewood Cliffs, New Jersey: Prentice-Hall.

Smith, Carlota S. 1961. A class of complex modifiers in English. Language 37: 11-54.

_____. 1964. Determiners and relative clauses in a generative grammar of English. Language 40: 37-52.

SUSSEX, Roland. 1974. The deep structure of adjectives in noun phrases. Journal of Linguistics. 10: 111-131.

VAILLANT, Andre. 1942. L' Article en vieux slave. Revue des etudes slaves 20: 5-12.

ZALIZNJAK, A.A. 1967. O pokazateljax množestvennogo čisla v russkom sklonenii. In To honor Roman Jakobson. Vol. 3. The Hague: Mouton. pp. 2328-2333.

Studies in the Linguistic Sciences
Volume 8, Number 1, Spring 1978

ASPECTS OF IRAQI ARABIC VERBAL PHONOLOGY*

David Odden

In this paper, I shall investigate certain phonological processes of Iraqi Arabic, based on data from Erwin (1963), Woodhead (1967) and supplemented by data from a native speaker of Iraqi Arabic. Since the majority of productive and regular morphology is found in the verbal system, the majority of forms for this study will be verbal. The primary descriptive goal of this study is to account for the stem and affix alternations in the various verbal paradigms. I shall argue that imperfective verb stems with the superficial shape CCVC are to be derived from an underlying CVCVC pattern via a right-to-left iterative Syncope rule. It is noted that a similar rule exists in Tonkawa which must apply in a left-to-right manner; this fact refutes a claim made by Jensen and Stong-Jensen (1973) that the direction of iteration for a rule can be determined by universal principles. An arguement for the cyclic application of the Stress Assignment rule is advanced, and a significant difference in the interaction between Stress Assignment and Epenthesis in Iraqi and Palestinian Arabic is noted.

The Iraqi verb appears in a perfective, imperative and imperfective conjugation. An optional object pronoun suffix is available in all conjugations. The perfective, imperfective, object pronoun and imperative affixes are provided below.

(1)

	Perf. Subject	Imperf. Subj.	Pro. Obj.	Imperative
1s	-t	ʔa-	-ni	
1p	-na	ni-	-na	
2m	-t	ti-	-ak	ʔi-
2f	-ti	ti-...-iin	-ič	ʔi-...-i
2p	-tu	ti-...-uun	-kum	ʔi-...-u
3m	-Ø	yi-	-a	
3f	-at	ti-	-ha	
3p	-aw	yi-...-uun	-hum	

Morphemes preceded by a dash follow the verb root and those followed by a dash precede the root. The subject pronoun sets associated with perfective, imperfective and imperative verb stems when realized as suffixes precede the object suffixes.

1. Stress and Epenthesis

The Stress Assignment rule of this language is essentially identical to the rule found in Classical Arabic and other Eastern Arabic dialects. A final heavy syllable (i.e. VCC or V̄C) is stressed, otherwise a penultimate heavy syllable is stressed, and the antepenultimate syllable is stressed if the two preceding conditions are not met. The following examples illustrate this pattern.

(2)	9alée	'on him'	naádi	'club'	mádrasa	'school'
	yihmárr	'he blushes'	9iraáqi	'Iraqi'	šárika	'company'
	mariiḍ	'sick'	sallátha	'her basket'	mumáccila	'actress'

As this pattern represents the stress pattern of virtually every word in the language, I assume the following Stress Assignment rule.

(3) V → [+stress] / __$C_0((\check{V}C)\check{V}(C))$]

A rule which interacts with Stress Assignment is found in the grammar which lengthens a vowel before the object pronouns and the dative preposition -il when it is cliticized to the verb.

(4)	kitábna	'we wrote'	kitabnaáhum	'we wrote them'
	dirásti	'you f. studied'	dirastiína	'you f. studied us'
	nísa	'he forgot'	nisaáha	'he forgot her'
	dazzeétu	'you p. sent'	dazzeetuúla	'you p. sent to him'

The following rule is proposed here to account for this lengthening phenomenon, although the precise formulation of the rule is not critical for this study.

(5) V → [+long] / ___ + {Dative Clitic, Object Pronoun}

I shall now turn to a consideration of some phonological alternations at work in the perfective paradigm. The first set of forms suggests the existence of a Syncope rule deleting an unstressed vowel in a doubly open syllable (i.e. when in the environment VC___CV).

(6)	kítab	'he wrote'	xaábar	'he phoned'
	kítbat	'she wrote'	xaábrat	'she phoned'
	kítbaw	'they wrote'	xaábraw	'they phoned'
	kitábit	'you m. wrote'	xaabárit	'you m. phoned'
	kitábti	'you f. wrote'	xaabárti	'you f. phoned'
	kitábtu	'you p. wrote'	xaabártu	'you p. phoned'
	kitábit	'I wrote'	xaabárit	'I phoned'
	kitábna	'we wrote'	xaabárna	'we phoned'

Assuming the underlying stems to be kitab and xaabar, the following Syncope rule accounts for the stem alternations kitab~ kitb and xaabar~ xaabr. Examples such as kitábit and xaabárit motivate the [-stress] condition in the focal vowel and also indicate the necessity of restricting (7) from applying after word initial consonants. Examples will be provided in a later portion of this section showing the need to restrict this rule from applying after consonant clusters.

(7) $\underset{[-\text{stress}]}{V} \rightarrow \emptyset$ / VC___CV

Additional examples of this pervasive rule are provided in (8), with stems followed by -a marking feminine, -een marking dual, -a marking one of a collective and -iič marking 'your f.'.

(8)	bágal	'mule'	bágla	'female mule'
	ʕálib	'dog'	ʕálba	'bitch'
	ṣáxal	'goat'	ṣáxla	'female goat'
	dáris	'lesson'	dárseen	'two lessons'
	gábur	'grave'	gábreen	'two graves'
	kaatib	'clerk'	kaatbeen	'two clerks'
	báhar	'sea'	báhreen	'two seas'
	táxat	'bench'	táxteen	'two benches'
	rutab	'date'	rutba	'a date'
	šaʕar	'hair'	šaʕra	'a hair'
	simač	'fish'	simča	'a fish'
	waʕad	'promise'	waʕdič	'your f. promise'
	kutub	'books'	kutbič	'your f. books'

Returning to the paradigms of (6), it will be observed that the 1st and 2nd singular forms kitábit and xaabárit have an anomalous phonetic shape in terms of stress placement, since one would expect *kítbit and *xáabrit on the analogy of 3 fem. kítbat, xáabrat, if the relevant verbal endings were taken to be underlying -it. The anomaly has an explanation, however, if we assume the underlying form of the ending to be -t and that the following Epenthesis rule is involved, ordered crucially after Stress Assignment (3).

(9) $\emptyset \rightarrow i \;/\; C\_\_\_C\begin{Bmatrix} C \\ \# \end{Bmatrix}$

This rule is supported by the fact that Iraqi has no final consonant clusters or internal clusters longer than two consonants. By this analysis, kitábit has the underlying representation kitabt. Stress falls on the final vowel according to (3), thus kitábt, and subsequent application of Epenthesis (9) derives correct kitábit. If, on the other hand, we assume the relevant terminations to be -it, there is no natural explanation possible.

There is additional evidence supporting the assumption that the vowel found in kitábit is epenthetic, beyond the explanation this assumption provides for Stress Assignment. A rule (or set of rules) inserts ee between stem and consonant initial suffix if the stem ends in a vowel or geminated consonant, and optionally if the stem contains a medial glide (which is deleted) and a derivational prefix (such as in nlaam 'be blamed', from nlíwam).

(10)	nísa	'he forgot'	dázz	'he sent'	nlaám	'he was blamed'
	nísat	'she '	dázzat	'she '	nlaámat	'she '
	nísaw	'they '	dázzaw	'they '	nlaámaw	'they '
	niseét	'you m. '	dazzeét	'you m. '	nlaameét	'you m. '
	niseéti	'you f. '	dazzeéti	'you f. '	nlaameéti	'you f. '
	niseétu	'you pl. '	dazzeétu	'you pl. '	nlaameétu	'you pl. '
	niseét	'I '	dazzeét	'I '	nlaameét	'I '
	niseéna	'we '	dazzeéna	'we '	nlaameéna	'we '

These three types of stems have in common the fact that they are in a sense "shorter" that the common CVCVC stem, lacking on the surface the final vowel or medial or final consonant. They thus may be characterized

as being stems of the form $C_oVC_o(V)$. Whether this is the correct designation for such stems or not is not crucial in this discussion; the important point is that the application of the ee-Epenthesis rule is conditioned to the right by a consonant initial suffix.

(11) $\emptyset \rightarrow$ ee / $[C_oVC_o(V)$ ___ +C

The fact that nisé́t exhibits epenthetic ee whereas nísat does not is explainable by the hypothesis that underlying nisa+at does not satisfy the conditions of the epenthesis rule, whereas underlying nisa+t for 'I forgot' does. This contrast is brought out in the derivations below (elision of a will be considered later).

(12)	nisa+at	nisa+t
Epenthesis		nisaeet
Elision	nisat	niseet
Stress	nísat	niseét

Even though -t sometimes has the vowel i, -t nevertheless behaves like a consonant initial suffix in inducing ee-Epenthesis.

2. Some Elision Processes

Verb stems having a medial consonant cluster do not undergo Syncope when the stem precedes the endings -at and -aw, as suggested in the formulation of Syncope (7). Instead, unstressed a is reduced to i.

(13)				
	staxdam	'he hired'	staxdimat	'she hired'
	?abdaθ	'he made'	?abdiθat	'she made'
	sta9mal	'he used'	sta9milat	'she used'
	?anjaz	'he did'	?anjizat	'she did'

The full perfective paradigm of staxdam is provided below: each verb of the type found in (13) conjugates in a similar fashion.

(14)				
	staxdam	'he used'	staxdamit	'I used'
	staxdimat	'she '	staxdamna	'we '
	staxdamit	'you m. '	staxdamtu	'you pl'
	staxdamti	'you f. '	staxdimaw	'they '

The following rule accounts for these alternations.

15) $a \Rightarrow i$ / $[+seg]C$ ___ CV +verb
$[-stress]$

The patterns of reduction are limited to verbs, since nouns, for example mádrasa 'school', mármara 'pies of marble', sfarjala 'quince', exhibit no reduction of a to i. The -stress condition is required to prevent reduction of a in staxdámit. The lefthand context [+seg]C ___ is required to prevent reduction of the initial vowel on ?abáddil 'I change'.
It will be shown below that reduction also applies in the context V:C___CV.

It might be assumed that Reduction applies to underlying kítabat yielding kítibat, and a revised version of Syncope (7) would apply which deletes only the high vowels i and u in the relevant environment. However, Reduction does not apply in nouns, yet unstressed a deletes in ʂáxla,'female goat', from ʂáxala. Therefore, Syncope is correctly formulated to delete all unstressed vowels.

We may now return to additional examples of Syncope in order to show restrictions which must be placed on that rule. Syncope also applies in the imperfective and imperative, assuming that the imperfective stem has the shape CVCVC (I shall argue for this position below). The data in (16) show that Syncope must apply from right to left; the underlying stems (containing the vocalism of the imperfective) are libas 'wear', tila9 'go out', kitib 'write' and suqut 'fail'.

(16)	yílbas	'he wears'	ʔilíbsi	'wear (f)!'
	yítla9	'he goes out'	yitil9uun	'they go out'
	tíktib	'you write'	tikítba	'you write it'
	yúsqut	'he fails'	yusuqtuún	'they fail'

A full paradigm of libas is provided below.

(17)	yílbas	'he wears'	ʔálbas	'I wear'
	tílbas	'she wears'	nílbas	'we wear'
	tílbas	'you m wear'	tilibsuún	'you pl. wear'
	tilibsíin	'you f wear'	yilibsuún	'they wear'

Within the imperfective, verbs of the shape CVCVC (such as libas) have the form CCVC when followed by a consonant initial suffix or no suffix, and have the form CVCC when followed by a vowel initial suffix. Thus, yilibsuún derives from yilibasuún via application of Syncope (7) to the rightmost vowel in a doubly open syllable, a. The consonant cluster in the output of this rule created by application of Apocope prevents reapplication of Syncope to i. If Syncope were to apply from the left, incorrect *yilbasuún would result.

Jensen and Stong-Jensen (1973) have argued that the direction of application of rules can be predicted by universal principles. One principle is that "accent" rules such as Syncope apply in the direction that minimizes application. Phelps (1975) has shown that their principles do not predict the direction of a rule such as Syncope, where both directions of application are equally minimizing. Phelps shows that Tonkawa has the following Syncope rule.

$$(18)\quad \underset{[+\text{stem}]}{V} \rightarrow \emptyset \;/\; VC((?)V)\_\_\_\; C\;V_{[-\text{suffix}]}$$

This rule must apply from left to right in Tonkawa. However, I have shown that Syncope in Iraqi must apply from right to left. Since the rule in Tonkawa and the rule in Iraqi are fundamentally the same rule and differ in direction of their application, the claim that directionality can be predicted does not appear to be wholely tenable.

One might wish to argue for a different analysis of the derivation of yilbas, as deriving from an underlying stem lbas employed in the imperfective and derive yilibsuún from yilbasuún via some rule of metathesis which also changes vowel quality. One might then propose the following Metathesis rule with Raising.

(19) V+CCVC+V ⇒ V+C V CC+V
[+hi]

This rule directly converts underlying yilbasuún to yilibsuún. The placement of morpheme boundaries is crucial in this rule. For example, if the righthand + is not specified, the rule incorrectly applies to the stem qtirih, which contains four radical consonants, in yi+qtirih 'he suggests', generating incorrect *yiqitrih. If the lefthand boundary is not specified, underlying staxdamat incorrectly becomes *staxidmat rather than staxdimat 'she used'.

Under the Syncope analysis, the independently motivated Syncope rule is employed to account for the stem alternations within the imperfect. Under the Metathesis-Raising analysis, an additional Metathesis rule is required. Furthermore, a special canonical shape of the stem, CCVC, is required in the imperfective, since it is assumed that the surface shape of the imperfective libs found in yilibsuún 'they wear' derives directly from the underlying stem lbas. The grammar must contain some statement describing the occurrence of CCVC stems in the imperfective, since not all stems have the shape CCVC on the surface. Note the following verbs which do not have the shape CCVC in the imperfect: in each case, the imperfect and perfect are formed from the same stem.

(20)	?á+qtirih	'I suggest'
	?a+staxdam	'I employ'
	?a+baddil	'I change'
	?a+haájaj	'I argue'
	?a+sírr	'I insist'
	?a+suuf	'I see'
	?á+ntirid	'I am fired'

Thus, it is only stems of the shape CVCVC in the perfect which have the shape CCVC in the imperfect. One can capture this relationship by the following rule:

(21) V ⇒ Ø / V+C__CV Imperfective

This rule does not claim that a phonological vowel deletion process converts the perfective stem into the imperfective, merely that a relation exists between these two stems.

The statement capturing the relation between imperfect and perfect stems under the Metathesis-Raising analysis is nearly identical to the independently needed Syncope rule. The Metathesis-Raising analysis is therefore to be rejected because it contains an unneeded metathesis rule to handle the data in (16), Syncope to handle the data in (6) and (8), and a stem formation rule for the imperfectives of CVCVC perfects which is virtually identical to Syncope. The analysis advanced here assumes a single stem to underlie the perfect and imperfect and a single Syncope rule.

One condition which must be placed on Syncope is that a vowel is elided if the lefthand context contains a geminated consonant. To that extent, geminate consonants do not behave like clusters.

(22)	báddal	'he changed'	báddlat	'she changed'
	ybáddil	'he changes'	ybáddluún	'they change'
	?ájjal	'he delayed'	?ájjlat	'she delayed'
	y?ájjil	'he delays'	y?ajjluún	'they delay'

The complete perfective and imperfective paradigm of báddal 'change' is provided below.

(23)	?abáddil	ls. imperf.	baddálit	ls.	perf.
	nbáddil	1p. "	baddálna	1p.	"
	tbáddil	2m. "	baddálit	2m.	"
	tbaddlíin	2f. "	baddálti	2f.	"
	tbaddluún	2p. "	baddáltu	2p.	"
	ybáddil	3m. "	báddal	3m.	"
	tbáddil	3f. "	báddlat	3f.	"
	ybaddluún	3p. "	báddlaw	3p.	"

On the other hand, if the righthand context contains a geminate consonant, an unstressed vowel is not elided.

(24)	?axallíha	'I let her'
	?abaddálkum	'I change you p.'
	kitabilkumiyyaá	'he wrote it to you p.'
	yistimirruún	'they continue'
	yišma?izzuún	'they are disgusted'

While deletion of vowels after geminate consonants in (22) seems to indicate that geminate consonants behave like single consonants and should therefore be represented as single consonants with the feature +long , the retention of vowels before geminate consonants in (24) suggests that geminates must be treated as clusters of consonants. Geminate consonants in Iraqi are thus an example of the dual behavior of geminates in language in general; in this example, the very same segments behave in different ways, depending on which side of the focus they fall on. The following revision of Syncope reflects this fact, admittedly in a rather unsatisfactory manner.

(25) $$\begin{matrix} V \\ [\text{-stress}] \end{matrix} \Rightarrow \emptyset \;/\; V(C_i)C_i \_\_\_\_ CV$$

The second restriction on Syncope is that it cannot bring two identical consonants together. The unstressed vowel between identical consonants remains unelided and is reduced to i by Reduction (15).

(26)	haájaj	'he argued'	haájijat	'she argued'
	?aθθaθ	'he furnished'	?aθθiθat	'she furnished'
	traášaš	'he moved'	traášišat	'she moved'
	báddad	'he wasted'	báddidat	'she wasted'
	txaššaš	'he specialized'	txáššišat	'she specialized'

This restriction is contained in (27).

(27) $V \rightarrow \emptyset \;/\; V(C_i)C_i\_\_C_jV \qquad i \neq j$
$[-\text{stress}]$

Again, this is a rather unsatisfactory representation of the restriction, although I believe it correctly represents a restriction on Syncope rather than auxiliary epenthesis rules. There is no evidence that the stem vowel ever elides in the relevant environments in (26), and if Syncope were to have applied to ḥaáj+at, yielding ḥaájj+at, it would merge that form in shape with forms such as dázzat 'she sent' and saádda 'his having closed', so no epenthesis rule can rescue the derivation of ḥaájijat.

A derivation of representative forms is provided below. Note that reduction of a in ḥaájijat provides an example of reduction after a long vowel.

(28)					
	baddad+at	baddal+at	staxdam+at	ḥaajaj+at
Stress	báddad+at	báddal+at	stáxdam+at	ḥaájaj+at
Syncope	NA	báddlat	NA	NA
Reduction	báddidat		stáxdimat	ḥaájijat

The next rule I propose deletes i from the imperfective prefixes when in an open syllable and followed by a closed syllable (i.e. VCC of VVC). This rule thus deletes the vowels of the imperfective prefixes ti-, yi-, and ni-, but not in ?a-.

(29)						
	ybáddil	'he changes'	tbáddil	'she changes'	nbáddil	'we change'
	ysaáfir	'he journeys'	tsaáfir	'she journeys'	nsaáfir	'we journey'
	ysiid	'he closes'	tsiid	'she closes'	nsiid	'we close'
	ynaám	'he sleeps'	tnaám	'she sleeps'	nnaám	'we sleep'
	ydeewur	'he turns'	tdeewur	'she turns'	ndeewur	'we turn'
	ybárhin	'he proves'	tbárhin	'she proves'	nbárhin	'we prove'

Elision of i in these prefixes is to be contrasted with retention of a in the first person singular forms of (30).

(30)				
	?abáddil	'I change'	?asaáfir	'I journey'
	?asidd	'I close'	?anaám	'I sleep'
	?adeéwur	'I turn'	?abárhin	'I prove'

The following Prefix Elision rule accounts for the reduction of prefixes in (29).

(31) $i \rightarrow \emptyset \;/\; \#C\_\_ + CV\begin{Bmatrix} V \\ C \end{Bmatrix}C$

The disjunction $\{V,C\}$ represents a strong syllable; this notation is unsatisfacto for expressing the notion 'strong syllable', but no more enlightening notation is available. The condition that i must itself be in an open syllable is shown by nonelision in the following examples where i is followed by two consonant

(32)				
	yitxárraj	'he graduates'	yitraáhan	'he bets'
	yitxáṣṣaṣ	'he specializes'	yiqtirih	'he suggests'
	yinkısir	'he is broken'	yistáxdim	'he employs'

The presence of the morpheme boundary in the rule is motivated by the fact that kitábti 'you f. wrote' retains initial i where no + boundary is present.

I shall show that verbs such as naám 'sleep' derive from the structure CVGVC via a Glide Deletion rule. Since the vowel of the imperfective prefix is elided in ynaám 'he sleeps' by Prefix Elision which requires the context CVVC, Glide Elision must precede Prefix Elision so that the requisite VV cluster is created. Prefix Elision must precede Syncope, however, since a stem of the shape CVCC derived from CVCVC via Syncope does not condition (31); thus yikitbuún 'they write' does not show Prefix Elision.

3. Vowel Sequences

This section will be concerned with the question of how sequences of vowels are modified when created by deletion of intervocalic glides or by morphological juxtaposition. Verbs having initial glides that appear phonetically in the perfective when initial exhibit deletion of that glide intervocalically after the imperfective prefixes. The prefix vowel and following glide appear to combine as follows: i+w → oo , i+y → ee , a+w → oo , a+y → ee . In addition to these alternations which hold true for all glide initial verbs, the two verbs ?axad 'take' and ?akal 'eat' show apparent combination of the glottal stop and prefix vowel such that a+? → aa , i+? → aa . Glides appear phonetically in the perfective forms yi?as 'he despaired', yibas 'he dried', wugaf 'he stopped', wusal 'he arrived'.

(33)	yee?as	'he despairs'	yeebas	'he dries'
	?ee?as	'I despair'	?eebas	'I dry'
	tee?suun	'you p. despair'	teebsuún	'you p. dry'
	yoogaf	'he stops'	yoosal	'he arrives'
	?oogaf	'I stop'	?oosal	'I arrive'
	toogfuun	'you p. stop'	toosluun	'you p. arrive'
	yaáxad	'he takes'	yaákal	'he eats'
	?aáxad	'I take'	?aákal	'I eat'
	taaxsuún	'you p. take'	taakluun	'you p. eat'

The stems ?akal and ?axad are peculiar in two ways. The first is that a glottal stop is deleted intervocalically in these forms, whereas all other verbs with glottal stop retain that glottal stop; thus, ?ilaf 'he was tame', yi?laf 'he is tame', yi?ilfuún 'they are tame'. Secondly, only these two verbs have the vocalic pattern CaCaC in the perfect (c.f. (6)), whereas all other CVCVC and CVCV verbs have CiCaC or CuCaC in the perfect. These two verbs also appear to have CaCaC as their vocalism in the imperfect, judging from the forms ?aáxad, yaáxad, whereas all other CVCVC verbs select either i or u as their initial stem vowel. If we assume these two stems to be axad and akal rather than ?axad and ?akal, with initial vowels functioning in the place of initial consonants in terms of canonical patterning, the divergence of these two stems from all other stems will be provided with a unified explanation.

We must now explain the remaining contraction problems. Two analyses appear to be at first glance plausible. In the first analysis, the sequences of glides and vowels are contracted by a single rule into a long vowel having features of the glide. In the second analysis, the analysis to be argued for here, features of the glide are present on the stem vowel, glides are deleted and the resulting vowel sequences will be appropriately contracted. The first

analysis requires the following rule.

(34) $V + \begin{bmatrix} -syl \\ -cons \\ +son \end{bmatrix} V \Rightarrow \emptyset \begin{bmatrix} +syl \\ -hi \\ +long \end{bmatrix} \emptyset$

This analysis however fails to account for yaáxuḁ which, deriving from underlying yi+axuḁ, does not contain the requisite glide for application of (34). Additionally, the glide-contraction analysis does not account for deletion of stem medial glides, as will be discussed below.

The vowel contraction solution we adopt here assumes that the high vowel after the glide in underlying yi+wugaf and yi+yi?as must be identical to the glide, and employs glide deletion and vowel contraction rules. Thus, yoógaf derives from yi+wugaf via the stage yi+ugaf. An identity constraint on high vowels after glides can be argued to be independently necessary for the imperfective. Glide Elision is governed by the following rule.

(35) $G \Rightarrow \emptyset \; / \; V \_\_\_\_ V$
$[-long]$

In addition to Glide Elision, a Vowel Contraction rule is necessary to complete the derivation. It will be observed that the resultant vowel is in all instances a lengthened and lowered stem vowel. This generalization is stated in the following rule.

(36) $V + V_i \Rightarrow V_i$
$[+stem] \quad [+long, +hi]$

The following derivations illustrate application of these rules.

(37)

	yi+akul	yi+wugaf	?a+yibas
Glide Elision		ø	ø
Contraction	yaákul	yoógaf	?eébas

Verbs which contain a glide as the second radical consonant retain that glide phonetically if the prededing vowel is long or the glide itself is long. A stem alternation exists in case the glide is deleted so that CaaC appears in the perfect before vowel initial endings or when the root is followed by ø and CiC or CuC appears in the perfective when followed by a C initial ending. Compare the perfective conjugations of ẍaáf 'see' and baág 'steal'.

(38)

ẍaáf	'he saw'	baág	'he stole'
ẍaáfat	'she " '	baágat	'she " '
ẍaáfaw	'they" '	baágaw	'they " '
ẍífit	'you m.'	búgit	'you m. '
ẍífti	'you f.'	búgti	'you f. '
ẍíftu	'you p.'	búgtu	'you p. '
ẍífit	'I '	búgit	'I '
ẍífna	'we '	búgna	'we '

The choice of vowels in the shortened (preconsonantal) forms might be expected to depend on the nature of the underlying glide, based on the facts of Classical Arabic. However, this is not the case, as can be seen below, since the glide w can correspond to either u or i in preconsonantal forms. If the imperfective stems contains a high vowel rather than aa (as in xaaf 'fear'), that vowel is invariably dependent on the underlying glide.

(39)	sifit	'I saw'	sáwwaf	'make see'	ysúuf	'he sees'
	bi9it	'I sold'	báyya9	'make sell'	ybií9	'he sels'
	ríhit	'I went'	ráwwah	'make go'	yruúh	'he goes'
	zírit	'I visited'	záwwar	'guide'	yzuúr	'he visits'
	jíbit	'I brought'	jáyyab	'attend'	yjíib	'he brings'
	gúmit	'I got up'	gáwwam	'make get up'	yguúm	'he rises'
	búgit	'I stole'	báwwag	'steal hab.'	ybuúg	'he steals'
	xúfit	'I feared'	xáwwaf	'frighten'	yxaáf	'he fears'
	dínit	'I owed'	dáayan	'lend to'	ydíin	'he owes'

We notice that there are two forms which indicate the identity of the medial glide in the above stems. The glide itself appears when preceded by a long vowel or when itself long, and secondly, a cognate high vowel appears in the imperfective. In general, the final stem vowel of the imperfective can be either a, i or u. A remarkable fact about glide medial stems is that no verb has y as its medial glide and uu as its imperfective vocalism or w as its medial glide and ii as its imperfective vocalism. That is, we never find verbs like *sayyaf ~ *ysuuf or *mawwat ~ *ymiit. This fact suggests some constraint on the relation between glides and imperfective vowels. One such constraint might state that vowels after glides in the imperfect must be identical to the preceding glide. But stems such as yxaáf 'he fears' (xáwwaf 'frighten') and ybaát 'he spent the night' (báyyat 'put up for the night') show that a appears as a stem vowel freely. Thus at least some nonidenti vowels appear after glides in the imperfect. The correct constraint is therefore on high vowels, that any high vowel after a glide in the imperfect is identical to the preceding glide.

(40) $$V_{[+hi]} \Rightarrow \begin{bmatrix} \alpha\,\text{back} \\ \alpha\,\text{round} \end{bmatrix} \;/\; \underset{\_\_\_\_}{G \begin{bmatrix} \alpha\,\text{back} \\ \alpha\,\text{round} \end{bmatrix}} \qquad \text{imperfective}$$

This constraint explains why a can appear freely in glide medial stems in the imperfective but i and u are restricted. At the same time, this constraint explains why all stems with initial glides have identical vowels, on the basis of the more general fact that the initial vowel for all CVCVC stems must be high in Iraqi Arabic (with the exception of akal and axaḏ).

A further fact which supports the assumption of a constraint on high vowels after glides in the imperfective is that as far as I have determined, all stems with medial glides which are retained require a following high vowel to be identical to the glide; thus, y9áyyin 'he appoints', ydaáyin 'he lends' yfáwwur 'he boils', ydáwwum 'he continues', ydéewur 'he turns', yistájwub 'he questions'. I conclude that the highvowel constraint expressed in (40) is correct.

Turning to the mechanism for selection of CaaC versus CiC as shown in (38), I assume that the stems underlying these alternations are siwaf and buwag, which become šiaf and buag by application of Glide Elision. These stems are then shortened by deletion of the final vowel before consonant initial suffixes by the following rule.

(41) $V \rightarrow \emptyset$ / V___CC

Notice that this rule has applied to šifit 'I saw', although the required consonant cluster is not found on the surface. If we assume that the 1st singular and 2nd m. suffixes are underlying -t as we have assumed earlier for other reasons, and assume a rule of Epenthesis such as (9), this anomaly is automatically explained. These forms have the intermediate representati šiaft, which becomes šift by application of (41) and šifit by application of Epenthesis. In assuming underlying -t rather than -it for the 1st singular and 2nd m. perfective suffixes, we have explained three otherwise unexplainable and unrelated facts, viz. Stress Assignment, ee-epenthesis and Shortening.

Having accounted for the mechanism which shortens glide medial stems, we turn now to the mechanism for generating the allomorphs with long aa, which will be derived from ua or ia. It will be noticed that the long vowel resulting from the vocalic cluster is essentially the righthand vowel, as was seen to be the case with glide initial and the two vowel initial verb stems in the imperfective. A rule assimilating ua and ia to aa already exists in our analysis, Contraction (36). Unfortunately, that rule as formulated is not applicable to the relevant vowel sequences in glide medial verbs, since (36) requires a morpheme boundary to intervene between the vowels of the cluster. If this boundary were removed from the rule, the assimilation would correctly apply to šiaf, yielding šaáf. The revised rule would then lower all sequences of two vowels within a stem, and should entail that no stem can have two high vowels consecutively, but would predict long midvowels whenever they should arise. This is clearly false, cf. yšuúf 'he sees', yjiíb 'he brings'. These forms and the forms in (39) unambiguously show that no vowel lowering is active within glide medial stems.

It is therefore necessary to separate the lengthening and lowering features of (36) into two rules with different domains of application so that these data can be properly accounted for. Since Lowering only applies between a prefix and a stem initial vowel, Lowering is restricted to that environment. The specification that the focal vowel is a stem vowel is necessary so that suffixes such as the feminine imperative -i do not undergo Lowering after vowel final verb stems, since ?i+bna+i becomes ?íbni 'build(f)!', not *?ibne (this will become clear immediately following).

(42) $\underset{[+\text{stem}]}{V} \rightarrow [-\text{high}]$ / V+___

The lengthening rule on the other hand has a wider domain of application, applying within glide medial stems as well as glide initial stems. It is not a general condition of the language that whenever two short vowels come together, a long vowel results. Therefore, lengthening must be stated as a specific rule and is not a metacondition on rules in Iraqi.

(43) $V \rightarrow [+long] / V\_\_\_\_$
$[+stem]$

To review this analysis, I have assumed that vowel sequences are created by Glide Deletion and by prefixing vowel final affixes to the two vowel initial stems. A constraint on high vowels after glides disallows *yu and *wi in the imperfective stem. The observed relations between deleted glide and resultant long vowel are analyzed as the result of this identity constraint on high vowels and specific rules to lengthen the second vowel of a VV sequence, plus a rule lowering the stem initial high vowels. A derivation of šaáf and yeé?as is provided below.

(44)			
		šiwaf	yi+yi?as
	Glide Elision	Ø	Ø
	Lengthening	aa	ii
	Lowering	NA	ee
	others	šaáf	yeé?as

The final rule needed to complete the derivation of aa from ia is a vowel truncation rule. This rule is necessary to explain facts other than those discussed immediately above. Whenever two vowels come together by suffixation, one of the vowels is deleted. This process has been alluded to in the discussion of ee-epenthesis; there, we noted that when a vowel final verb stem such as bína 'build' is followed by a vowel initial suffix or ee is inserted by ee- Epenthesis (11), the stem vowel is deleted. In the following examples, the lefthand example shows the stem vowel followed by no suffix, while the righthand example consists of stem plus affix.

(45)				
	bína	'he built'	bínat	'she built
	yíbni	'he builds'	yibnuún	'they build'
	?íqra	'read(m)!'	?íqri	'read(f)!'
	yínmu	'he grows'	tinmiín	'you (f) grow'

These data suggest the following rule.

(46) $V \rightarrow Ø / \_\_\_ +V$

When we bring into consideration the effect of object pronouns added to verbs, we require an extension of this rule. Recall that before object pronoun suffixes, a vowel is lengthened by Lengthening (5), as in kitabnaáhum 'we wrote them', from kitabna+hum. If a vowel initial object suffix is added to a vowel final stem, the stem vowel is lengthened by this rule and the suffix vowel is deleted. Thus, the vowel of -iš 'you (f)', -ak 'you (m) ' and -a 'him' is deleted in the examples below.

(47)				
	dirábna	'we hit'	dirabnaáč	'we hit you(f)'
	níša	'he forgot'	nisaáš	'he forgot you (f)
	tínmu	'he grows'	tinmuú	'he grows it'
	qíra	'he reads'	qiraák	'he reads you(m)

We therefore see that a short vowel deletes before or after any long vowel, and before another short vowel. Assuming that mirror image rules abbreviate rules in the order {/____X , /X_____} , these deletions can be accounted for by modifying (46) as follows.

(48) $V \rightarrow \emptyset \;/\; \underset{[-long]}{\quad} V$

4. Stress and the Cycle

I shall now turn to a consideration of the dative clitic -il and the pronominal link -yya which occurs underlyingly between the dative clitic and the direct object pronoun on a verb. There are two logical possibilities for the underlying form of each of these morphemes, given the existence of Syncope and Epenthesis: these are -il, -l, -iyya and -yya. I shall argue that the choice of -il and -yya as the underlying forms is the only selection which accounts for critical data. The vowel i appears in -il when that suffix stands before a consonant initial suffix, and is not found before vowel initial suffixes.

(49) kitab+íl+na 'he wrote to us' kitáb+l+iš 'he wrote to you(f)'
dazz+íl+na 'he sent to us' dázz+l+iš 'he sent to you(f)'

Under the analysis with underlying -il, the vowel is deleted by Syncope (which, it will be recalled, can delete an unstressed vowel after a geminated consonant, as indicated by (27)). Thus, kitáblič derives from kitábilič via Syncope.

The suffix -yya selects the vowel i after a consonant, and selects no vowel after a vowel, as indicated below.

(50) kitabilnayyaá 'he wrote it for us' kitabilkumiyyaá 'he wrote it for yo
dazzeetilnayyaá 'you sent it to us' dazzeetilhumiyyaá 'I sent it o them'

It will be noticed that the vowel of the indirect object suffix na in kitabilnayyaá does not become lengthened by Lengthening (5). This fact was anticipated in the formulation of that rule, since the pronoun linking morpheme -yya was not included in the set of suffixes which condition Lengthening. The fact that the preceding a vowel of na is not lengthened and yet not elided in favor of i clearly shows that kitabílnayyaá does not have the intermediate structure kitabilna+iyyaá, since according to (48), the lefthand vowel of a short vowel cluster elides (as in ?íqri 'read(f)!', from ?iqra+i). Assuming the underlying form -iyya for this suffix results in incorrect *kitabilniyyaá. Only the underlying form -yya explains the retention of the final object suffix vowels, by deriving the vowel of iyya through Epenthesis, whenever that vowel appears phonetically.

Having extablished the underlying form of -yya, we can now turn to forms which show that the underlying form of the dative clitic must contain the vowel i, and which also show the ordering relation between Epenthesis and Syncope (specifically, Epenthesis precedes Syncope). One argument for assigning the dative clitic the underlying representation -il is the fact that that vowel appears everywhere when not cliticized to the verb; thus ?ílha 'to her', ?íla 'to him'. When -il is followed by one of the vowel initial object suffixes -iš or -ak which are in turn followed by iyyaa, the vowel of those object suffixes is elided.

(51) kitabilkiyyaá 'he wrote it to you (m)'
kitabilčiyyaá 'he wrote it to you (f)'
dazzeetilkiyyaá 'I sent it to you (m)'

Assuming the indirect object clitic to be -il and assuming that Epenthesis feeds Syncope, these forms are derived in the following manner.

(52)

	kitabilakyyaá	kitabiliV́yyaá
Epenthesis	kitabilakiyyaá	kitabiliciyyaa
Syncope	kitabilkiyyaá	kitabliciyyaa

On the other hand, if we assume the dative clitic to have the underlying representation -l, or if we assume that Syncope precedes Epenthesis, incorrect results follow.

(53)

	kitabilikyyaá	kitablikyyaá	
Sync.	NA	kitablikiyyaá	Epenth.
Epenth.	kitabilikiyyaá	NA	Sync.
	*kitabilikiyyaá	*kitablikiyyaá	

The only analysis consistent with elision of i and a in -iC and -ak is that Epenthesis precedes Syncope and that the dative clitic has the underlying form -il.

At this point, we have seen the relevant rules of Iraqi Arabic, so we may now turn to a set of facts arguing that stress must be assigned cyclically. It will be recalled that I have shown that Stress Assignment applies before Epenthesis, since Stress Assignment applies to the abstract representation kitabt rather than derived kitabit. Thus, an epenthetic vowel is not present when stress is assigned. Now consider the following examples in terms of the ordering of Epenthesis and Stress Assignment, and also in terms of application of Syncope. These examples contain the suffix(es) t which receives its i through application of Epenthesis.

(54)

dirásit	'I studied'	dirasítha	'I studied her'
9irafit	'you knew'	9irafítha	'you knew her'

We notice two unusual facts about Verb + Object forms on the righthand side. Epenthetic i is stressed in a syllable closed by the person suffix and object suffix -ha. Furthermore, the stem final unstressed vowel in a doubly open syllable is not elided by Syncope. Since the epenthetic vowel is stressed, it must be present when stress is assigned, which means that Epenthesis must precede Stress Assignment. But we have already shown that Stress Assignment precedes Epenthesis and that the epenthetic vowel cannot be present when stress is assigned. Thus, an ordering paradox results.

The ordering paradox dissolves if Stress Assignment is applied cyclically. Bracketing dirasítha appropriately (ie. [[dirast]ha]), cyclic application of Stress Assignment allows that rule to both preced and follow Epenthesis. In addition, if stress is assigned to the final stem vowel in the course of the derivation as predicted by the cyclic hypothesis and later reassigned to the epenthetic vowel, presence of reduced stress on the final stem vowel prevents that vowels elision. The cyclic derivation of dirasítha is provided below.

(55)

	[[dirast] ha]
Stress	1
Epenthesis	[dirásitha]
Stress	[dirəsítha]
Syncope	NA

The cyclic application of Stress Assignment in Iraqi thus allows two anomalies to be related and explained without resorting to adhoc conditions on rules.

A different argument for cyclic application of Stress Assignment in Palestinian Arabic based on essentially different phenomena was provided by Brame (1974). A crucial difference between these two dialects exists in terms of their treatment of Stress and Epenthesis. Whereas in Iraqi an epenthetic vowel is stressable on a second cycle, as in dirasítha, in Palestinian Arabic, the epenthetic i is never stressable. Thus, Palestinian has darásit 'I studied', darásitha 'I studied her'. Nevertheless, these two dialects have the same Stress Assignment rule and Epenthesis rule. A language specific statement of Iraqi grammar specifies that Epenthesis may apply between the two cycles of Stress Assignment, whereas in Palestinian Arabic, Epenthesis does not intervene. Thus, the point should be made (and has not been made elsewhere) that part of the condition that a rule applies cyclically is that the set of rules which intervene between one cycle of a rule and the next cycle of a rule must be specified in the grammar. It is a question for future research how three cyclic applications of a rule are to be handled and what the interactions between two cyclic rules might be.

NOTES

*My thanks to Mike Kenstowicz and Chuck Kisseberth for having read and commented on an earlier draft of this paper. My thanks also to Saad Ahmad, a native of Baghdad, for providing crucial material for this study.

REFERENCES

BRAME,M. 1974. The cycle in phonology: Stress in Palestinian, Maltese and Spanish. Linguistic Inquiry 5.39-60

ERWIN,W. 1963. A short reference grammar of Iraqi Arabic. Georgetown Arabic Series #4: Washington.

PHELPS,E 1975 Iteration and disjunctive domain in phonology. Linguistic Analysis 1.137-172

JENSEN,J and M. Stong-Jensen 1973. Ordering and directionality of iterative rules. Papers in Linguistics 6.66-90.

WOODHEAD,D. 1967 A dictionary of Iraqi Arabic: Iraqi-English. Georgetown Arabic Series #10: Washington.

Studies in the Linguistic Sciences
Volume 8, Number 1, Spring 1978

EXCEPTIONS AND RULE GOVERNMENT: THE CASE OF THE PASSIVE RULE IN HINDI

Rajeshwari Pandharipande

The goal of this paper is to examine two major hypotheses (Lakoff 1965 and Green 1974, 1976) about exceptions to syntactic rules and their role in defining the concept of rule-government. In this context the case of the exceptions to the passive rule in Hindi is discussed. It is argued that the exceptions to the passive rule in Hindi do not present true irregularity and thereby fail to support Lakoff's hypothesis which treats all exceptions as true irregularity. The case of the Hindi passive supports the hypothesis presented in Green (1974, 1976) in that the exceptions are regular and systematic. The passive rule in Hindi applies to a semantically specifiable class of verbs, i.e., verbs expressing a volitional act. Thus, the verbs which do not express a volitional act fail to undergo the rule and present apparent irregularity. Passive is a governed rule in Hindi in the sense that it admits a semantic class in its structural description. It is pointed out with illustrations that the above hypothesis is independently motivated and is needed in the language.

1.0 Introduction

A number of recent studies in linguistics (Lakoff 1965, Lakoff, 1968, Green 1974, Green 1976, and Green and Morgan 1976) have discussed the phenomenom of exceptions to syntactic rules and their role in defining the concept rule-government. Two major questions were discussed in this context: (i) does every case of exception to the rule present true irregularity? and (ii) should a rule be labelled as governed on the basis of the fact that it admits exceptions? The following hypotheses have been proposed: hypothesis I (Lakoff 1965) treats all cases of failure to undergo a rule whose structural condition is satisfied as lexical exceptions (true irregularities). No distinction is made between lexical exceptions and other kinds of items which for some general and independently motivated reasons fail to undergo the rule. All rules which have exceptions are claimed to be governed rules. In contrast to this, hypothesis II (Green 1974, Green 1976, and Green and Morgan 1976) makes a distinction between lexical exceptions and apparent irregularity. It is claimed that rules can apply to the semantically specifiable class of lexical items. In such cases, lexical items which do not belong to the appropriate semantic class fail to undergo the rule and produce apparent irregularity. Governed rules, according to this hypothesis, are rules whose structural description mentions semantic classes.

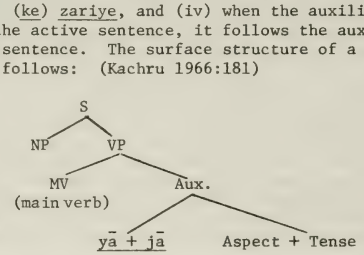
The purpose of this paper is to discuss the passive rule in Hindi (Passive hereafter) and investigate the nature of the verbs which fail to undergo the rule (exception-verbs hereafter). The major points of focus are the following: (i) Neither the traditional grammars nor hypothesis I provide any explanation for the "exceptional" behavior of the exception-verbs. Hypothesis I forces us to treat these exception-verbs as lexical exceptions. (ii) There is independently motivated evidence to show that these verbs are not true exceptions. They are systematic, i.e., they fall into a coherent

semantic class, namely, verbs expressing a non-volitional act. (iii) In order for the hypothesis to be descriptively adequate, it is necessary to treat Passive as a governed rule, in the sense proposed in hypothesis II. If it is claimed that Passive in Hindi applies to the verbs expressing a volitional act, it would explain in a straightforward fashion why certain verbs fail to undergo the rule. Also, it would correctly predict whether or not a borrowed verb in Hindi would undergo the rule. In the light of the above-mentioned facts, some observations about the semantics of passive sentences in Hindi are made. It is pointed out that passive sentences in Hindi express more than one meaning. However, they all share one thing in common, i.e., they all express a volitional act. Subject-less passive sentences in the present tense and habitual aspect express a convention and prescribes a particular mode of behavior, e.g., barõ se istarah zorse bolā nahĩ jātā 'it is not customary to talk loudly like this with elders.' This sentence thus expresses a convention, namely, one does not talk loudly with the elders, and thereby it is suggested that one should not talk loudly with the elders. It is observed that the same type of construction is found in the Brahmanical texts in Sanskrit literature. The purpose of these texts is to illustrate the ritualistic conventions and prescribe a particular mode of behavior.

This paper also includes discussion about the theoretical implications of this proposal (cf. 6.1).

1.1 Passive in Hindi

In this section, I will describe the passive rule in Hindi and present evidence to support the claim that Hindi has a passive rule. Passive in Hindi applies to both transitive and intransitive verbs. When a verb is passivized, certain changes take place in the sentence. These changes can be illustrated as follows: (i) the main verb of the active sentence is in its perfective form in the corresponding passive sentence; (ii) an auxiliary verb jānā 'to go' is attached to the main verb; (iii) the subject of the active sentence (ex-subject hereafter) is followed by one of the following postpositions; se, (ke) dwārā, (ke) zariye, and (iv) when the auxiliary verb honā 'to be' is present in the active sentence, it follows the auxiliary verb jānā 'to go' in the passive sentence. The surface structure of a passive sentence can be illustrated as follows: (Kachru 1966:181)



Compare the following non-passive sentences (1) and (2) with the passive sentences (1.a) and (2.a).

Intransitive verb calnā 'to leave'

(1) rām calā
Rām left.

(1.a) rām se calā gayā
ram by left went
(perfective)
It was left by Ram. (Ram left.)

Transitive verb gānā 'to sing'

(2) rām gāne gātā hɛ
ram songs sings aux
Ram sings songs.

(2.a) rām se gāne gāye jāte hẽ
ram by songs sang go aux
(perfective)
The songs are sung by Ram.

Notice that the verbs calnā 'to leave' (1) and gānā 'to sing' (2) are in their perfective form in (1.a) and (2.a) respectively. The subject of (1.a) and (2.a) is followed by the postposition se 'by.' Also, notice that the auxiliary hona 'to be' is present in (2) and is retained in (2.a), where it follows the auxiliary jānā 'to go.'

2.0 Verb-agreement

The verb-agreement is not the same in active and passive sentences. While the verb agrees with the subject in an active sentence, in a passive sentence it agrees with the object of the active sentence (ex-object hereafter). Consider the following examples:

(3) larkā citthiyā̃ paṛhtā hɛ
boy letters reads aux
3p.sg. 3p.plu. 3p.sg. 3p.sg.
mas. fem. mas.
The boy reads the letters.

(4) larke se citthiyā̃ paṛhī jātī hẽ
boy by letters read go aux
3p.sg. 3p.plu. 3p.plu. 3p.plu 3p.plu.
mas. fem. fem. fem.
The letters are read by the boy.

Notice that while the verb and aux. paṛhtā hɛ 'reads' agrees with the subject larkā 'boy' in (3), the verb paṛhī 'read' and both the auxiliaries agree with the ex-object citthiya 'letters' in (3.a).

An intransitive verb agrees with the subject in an active sentence. However, it remains in its unmarked form in a passive sentence. Consider the following examples:

(4) larke nahĩ soe
boys not slept
3p.plu 3p.plu.
mas. mas.
The boys did not sleep.

(5) larkõ se soyā nahĩ gayā
boys by slept not went
3p.plu. 3p.sg. 3p.sg.
mas. mas. mas.
It was not slept by the boys.
(The boys could not sleep.)

Notice that the verb soe 'slept' agrees with the subject larke 'boys' in (4) but it remains in its unmarked form in (4.a). The example (3.a) shows that when a transitive verb is passivized, if the ex-object is not followed by a postposition, the verb agrees with the ex-object in number and gender in all aspects and tenses. Notice that the subject in (3.a) is followed by the postposition se while the ex-object is not followed by any postposition. Now consider the following example of a passive sentence where both the ex-subject and the ex-object are followed by a postposition. In this case the verb does not agree with either of them.

(5) larkiyõ se zakhmī ādmiyõ ko nahĩ dekhā gayā
girls by injured people not saw went
3p.plu. 3p.plu.obj. 3p.sg. 3p.sg.
fem. mas. mas. mas.
The injured people were not seen by the girls.
(The girls could not bear to see the injured people.)

Notice that the ex-subject larkiyā 'girls' is followed by the postposition se. Also, the ex-object ādmī 'people' is followed by the postposition ko. The verb (in its perfective form) and the auxiliary dekhā gayā 'was seen' do not agree with either the ex-subject or ex-object.

The agreement rules for passive sentences can be stated as follows (for further discussion see Kachru 1966:182)

(i) When a transitive verb is passivized, if the ex-object is not followed by ko, the verb agrees with the ex-object in number and gender [recall (3.a)].

(ii) However, if the ex-object is also followed by a postposition ko, the verb remains in its unmarked form.

(iii) If an intransitive verb is passivized, it remains in its unmarked form.

The verb agreement in a passive sentence in Hindi abides by the axiom of the verb agreement in Hindi, namely, a verb does not agree with a noun which is followed by a postposition. Thus the case of the verb-agreement in the passive sentences is not different from the case of active sentences where the subject is followed by a postposition. In both cases the verb does not agree with the subject. Notice the following example of non-passive sentences:

(6) larkõ ne kām kiyā
boys ag. job did
3p.plu.ag. 3p.sg. 3p.sg.
mas. mas. mas.
The boys did the job.

(7) baccõ ko mã̃ yad ati hɛ̃
children mother memory comes
3p.plu. obj.m. 3p.sg. 3p.sg. 3p.sg.
mas. fem. fem.
The children remember the mother.

Notice that in both (6) and (7), the subject larke 'boys' in (6) and bacce 'children' in (7) is followed by a postposition (ne and ko respectively) and the verb is in its unmarked form in (6) and agrees with mā̃ 'mother' in (7).

2.1 Evidence for Passive: Reflexivization

In this section, I will present evidence to show that Hindi has a passive rule. The Hindi reflexive pronoun apnā 'self' refers to the subject of the Hindi sentence. Consider the following sentence:

(8) rām ne ṿsyām ko apnī jagah par bhejā
rām ag. ṿsyām obj.M. self's place on sent
Ram sent syam to his place.
Rām's
*ṿSyam's

Notice that the reflexive apnī 'self' refers to Rām (subject) and not ṿSyam (object). However, the following example shows that the reflexive in the corresponding passive sentence refers to the subject as well as the object.

(9) rām se ṿsyām ko apnī jagah par nahĩ̃ bhejā gayā
rām by ṿsyam obj.m. self's place on not sent went
ṿSyam was not sent by Rām to his place.
ẋSyam's
Ram's

The most straightforward way to account for the facts of reflexivization in Hindi is by analyzing (8) and (9) as being transformationally related. For, consider the alternative. If the above assumption is not made, sentences like (9) may lead one to believe that the reflexive in Hindi refers to subject as well as object. This is patently false, as sentence (8) makes clear. The fact is that the reflexive in Hindi refers to both subject and object if and only if the sentence is passive. In active sentences it refers only to the subject. This fact is easily accounted for if we treat (9) as being derived from (8) by the rule (Passive) which promotes the object to the subject position, then we can explain the behavior of apnī 'self's' in a straightforward fashion, i.e., examples (8) and (9) do not contradict the generalization namely, the reflexive apnī 'self's' always refers to the subject and not the object. It is because ṿSyam [object in (8)] is promoted to subject-position by the rule (Passive), that it behaves like the subject and therefore the reflexive apnī 'self's' can refer to ṿSyam [ex-object in (9)]. This evidence indicates the following:

(i) Hindi has a passive rule.

(ii) Passive is a relation-changing rule in Hindi

and (iii) Although (9) provides evidence for object-promotion, it does not provide any evidence for subject-demotion [recall (9) where the reflexive

continues to refer to rām (the ex-subject)] in a passive sentence. (For more discussion on the subject of passive sentences in Hindi, see Kachru et al 1976).

2.2 Evidence for Passive: Subject-deletion

The ex-subject is generally deleted in a passive sentence in Hindi. Thus Hindi has the following kind of passive sentences:

(11) ∅ kām kiyā gayā
work did went
The work was done.

If we do not consider this deletion as a consequence of the Passive, and thereby sentences like (11) as derived sentences, then we will have to treat these sentences as being identical to the other subject-less sentences like (12).

(12) kām huā
work happened
The work got done.

However, for the native speakers of Hindi, (11) and (12) are not semantically identical. For them the agent (subject) is implied in (11) while it is not in (12). This intuition of native speakers is correctly supported by the fact that when the agent and its agency is explicitly denied in (11) and (12) respectively, (11.a) presents contradiction while (12.a) does not.

(11.a) * kisīke na karne par bhī kām kiyā gayā
anybody's not doing inspite of work did went
The work was done without anybody's doing it.

(12.a) kisīke na karne par bhī kām ho gayā
anybody's not doing inspite of work happened
The work got done without anybody's doing it.

(11.a) indicates that the agent is implied in it while (12.a) indicates that it is not. The following assumptions explain the situation in a simple fashion.

(i) The underlying structure of (11) is (13), in which the agent is present. (11) is derived from (13) by a rule which generally deletes the ex-subject after the application of passive rule.

(13) X ne kām kiyā
X ag.M. work did
X did the work.

(ii) The underlying structure of (12) is (14) which is the same as its surface structure.

(14) kām huā
work happened
The work got done.

Notice that the subject is not present in the underlying structure of (12) whereas it is in the underlying structure of (11). Thus in order to differentiate the sentences like (11) from (12), it is necessary to assign to them two different underlying structures and derive (11) from (13) by the passive rule.

3.0 Exceptions

A number of verbs in Hindi fail to undergo the passive rule. The list of the exception-verbs consists of the verbs (intransitive) like dikhnā 'to seem,' khilnā 'to bloom,' murjhanā 'to whither,' galtī honā 'to make a mistake,' biknā 'to sell,' tūṭnā 'to break' and the verbs (transitive/intransitive) which require their subject to be in the dative/oblique case i.e., lagnā 'to feel,' pasand honā 'to like' (literally, liking happen), mālūm honā 'to know,' yād ānā 'to remember' (literally, memory come) and sirdard honā 'to have a headache' (literally, headache happen) etc.

The traditional grammars (Guru 1920, Vajpeyi 1958, Kellogg 1955), do not discuss these exceptions. Within the framework of hypothesis I (Lakoff 1968), these verbs will have to be treated as lexical exceptions (true irregularity) and will have to be marked (-Passive) in their lexical entry. Also, since the passive rule in Hindi allows these exceptions, it will be marked as a governed rule.

3.1 The "capabilitative" meaning

Before I present a semantically oriented explanation for the "exceptional" behavior of these verbs, it will not be inappropriate to see what a passive sentence means in Hindi. A passive sentence in Hindi, with the ex-subject present on the surface gnerally expresses the capacity of the subject to carryout the act expressed by the verb (Vajpeyi 1958, Kachru 1966). Consider the following sentences:

(15) laṛke se cɪtthiyɑ̃ nahɪ̃ likhī jatɪ̃
boy by letters not wrote go
The letters are not written by the boy.
(The boy is not capable of writing letters.)

However, whether or not a verb expresses capability of the subject, does not help us define the exceptions for the following reasons: (i) verbs which express capability of the subject, do not necessarily undergo Passive and (ii) not all passive sentences in Hindi express the capabilitative meaning (Kachru 1966:183-184). Passive sentences in Hindi cover a wide range of meaning. A descriptively adequate analysis of the "exceptional" behavior of the verbs should take into account all the meanings expressed by those sentences.

In what follows, I will present evidence to show that the capacity of the subject either to undergo or to carryout the act expressed by the verb is not crucial to decide whether or not a verb will undergo Passive. Consider the following sentences:

(16) phūl khiltā hɛ
flower blooms aux
The flower blooms.

(17) mujhe acchā lag rahā hɛ
to me good feel prog. aux
I am feeling fine.

Notice that the subject phūl 'flower' in (16) is capable of undergoing the act of blooming. Similarly, the subject mɛ̃ 'I' in (17) is capable of feeling (fine). However, when the verbs khilnā 'to bloom' and lagnā 'to feel' are passivized, (16.a) and (17.a) are ungrammatical.

(16.a) * phūl se khilā jātā hɛ
flower by bloomed goes aux.
It is bloomed by the flower.

(17.a) * mujh se acchā lagā jā rāhā hɛ
me by good felt go prog. aux.
It is felt good by me.

This evidence shows that the subject's capacity to carry out or to undergo the act expressed by the verb is not crucial to decide whether or not a verb will undergo Passive.

3.2 In this section I shall show that a passive sentence expresses the capabilitative meaning only when the ex-subject is present. When the ex-subject is deleted, passive sentences lose the capabilitative meaning (kachru 1966). Consider the following sentences:

(18) laṛke se citthiyā̃ nahĩ̃ likhī jātĩ̃
boy by letters not wrote go
(i) The letters are not written by the boy.
(ii) The boy is not capable of writing letters.

(19) ✓ citthiyā̃ nahĩ̃ likhī jātĩ̃
letters not wrote go
The letters are not written.

Notice that when the ex-subject is present, (18) expresses the capabilitative meaning. However, when the ex-subject laṛke se 'by the boy' is deleted, (19) loses the capabilitative meaning.

3.3 In (3.2) we have seen that subject-less passive sentences do not express the capabilitative meaning. Now I will present various meanings expressed by the subject-less passive sentences under different conditions. First, I will present subject-less passive sentences which express a convention, a custom or a tradition in the present tense and the habitual aspect. These sentences are used to prescribe a particular mode of behavior. Consider the following sentences:

(20) baṛõ se istarah zor se nahĩ̃ bolā jātā
elders with like this loudly not talked goes
(One does not talk) loudly like this with the elderly people.

(21) dīwālī par ghar ghar mɛ̃ dīp jalāe jāte hɛ̃
diwālī on house house in lamps lit go aux.
On the Dīwālī-day lamps are lit in every house.

(22) ām chīlkar nahĩ khāe jāte
mangoes having peeled not ate go
Mangoes are not eaten having been peeled.

Notice that the ex-subject is not present in (20), (21) and (22). (20) expresses a convention, namely, it is not customary to talk loudly with the elders and thereby suggests that it is not proper to talk loudly with the elders. (21) expresses the custom of lighting lamps of the Dīwālī-day (festival of lights, celebrated in India); and thereby suggests that one should light lamps on the Dīwālī-day. (22) expresses the convention that the mangoes are not peeled. (They are either cut into small pieces or their juice is squeezed out.) This sentence suggests that one should not peel mangoes. All the three sentences above [(20), (21) and (22)], carry an habitual aspect and present tense. These sentences lose their "prescriptive" meaning when their aspect or tense is changed. Consider the following examples:

(23) dīwālī par dīp jalāe gae
dīwālī on lamps lit went
(perfective) (past)
The lamps were lit on the Dīwālī-day.

(24) ām chīlkar nahĩ khāe jā rahe (hɛ̃)
mangoes having peeled not ate go prog. (aux)
The mangoes are not being eaten having been peeled.

Notice that (23) is in the perfective aspect and past tense, while (24) is in the progressive aspect and present tense. In both cases the "prescriptive" meaning, expressed by (21) and (22) is blocked. Neither (23) not (24) express any custom, tradition or convention, instead, they refer to a particular incident at a particular point in time. Thus in order for passive sentences to convey the "conventional" meaning, the role of the aspect and the tense seems to be crucial. The "capabilitative" meaning is also blocked in both (23) and (24).

This "prescriptive" usage of passive sentences in present tense and habitual aspect has precedent in the history of Indo-Aryan. The Brahmanical texts in Sanskrit make extensive use of subject-less passive sentences. The purpose of these texts is to interpret the Vedic texts in the context of the sacrificial rituals. These texts describe the sacrificial rituals and illustrate the duties of the priests and Yajamana (the host for whom the sacrifice is performed), etc. (For more discussion see Keith 1928 and Gonda 1975). A close examination of the data shows that subject-less passive sentences here are utilized to express conventions and thereby precribe particular modes of behavior. The following examples illustrate the situation.

(25) sāma gīyate
sacrificial prayer is sung
gai + yá (pass.) + present
The sama is sung. (Śatapatha Brahmana 5.2.46)

(26) āhutir hūyate
oblation is offered
hū + yá (pass.) + present
The oblation is offered. (Śatapatha Brāhmana 3.1.23,28)

(27) sa svāhetyevājuhot tasmāt u
he svaha thus saying offered therefore indeed
svāhetyeva huyate
svaha thus saying is offered
hū + yá (pass.) + present
He offered (the oblation) with (the words) "svāhā!".
Therefore (it is) offered with svāhā. (Śatapatha Brāhmana 2.2.4.6)

Notice that the subject is not present in (25)--(27). (25) expresses a convention, namely, the sama is sung. It also expresses a "prescriptive" meaning namely, the sama should be sung. (26) expresses a convention, i.e., the oblation is offered and thereby suggests that the oblation should be offered. In (27), the process of offering the oblation is described. The first half of the sentence describes an incident i.e., he offered the oblation, reciting the word "svāhā!" (an exclamation used in offering oblations to gods). The second half of the sentence describes the convention, 'therefore it is offered with the recitation of the word "svaha!".' This sentence conveys a "prescriptive" meaning, namely, the oblation should be offered with the recitation of the word "svāhā!". This "prescriptive" meaning conveyed by the passive sentences in the Brahmanas, is restricted to the present tense and the habitual aspect. The following sentences in the past tense do not convey this "prescriptive" meaning:

(28) tasmāc śaro nāma yad aśīryata
therefore arrow named because was broken
a(past M) + sri + past 3sg.
Therefore the designation arrow, because it was broken off.
(Śatapatha Brāhmana 2.4.1)

(29) sapta rsayah asrjyanta
seven śages were created
a(past) +srj + ya(passive) + past 3sg.
The seven sages were created.
(Śatapatha Brāhmana 8.4.3--6)

Note that (28) and (29) do not express a prescriptive meaning like (26) and (27). While (25)--(27) are in the present tense, (28) and (29) are in the past tense. In this section I have shown that:

(i) When the subject is deleted, the passive sentences in Hindi lose their "capabilitative" meaning.

(ii) In the present tense and the habitual aspect, the passive sentences express a convention and prescribe a particular mode of behavior;

(iii) The passive sentences in the Brahmanas behave similarly [For more discussion on the function of the passive sentences in the Śatapatha Brāhmana, refer to Pahdharipande:1976(unpublished).]

This similarity between the function of the passive sentences in the Brāhmanas and in Modern Hindi leads to a speculation namely, Hindi might have inherited this usage of passive sentences. However, in order to arrive at this conclusion, it is necessary to investigate the function of passive sentences in the literature written between the period of the Brahmanical texts and Modern Hindi. Prākṛt, Apabhraṃśa and Old Hindi mark three major stages of Indo-Aryan between the period of Śatapatha Brāhmana and Modern Hindi. If the investigation of the texts shows that the "prescriptive" function of passive sentences is retained through those stages in the development of Indo-Aryan, then it can be claimed that Modern Hindi inherited the "prescriptive" function of passive sentences. At this point this is an open question and needs further investigation.

3.4 In this section, I will show that in some "registers" (or 'functionally determined language-types' Halliday:1964) of Hindi[2] (i.e., the news paper register and the scientific register), the subject-less passive sentences express an "impersonal" meaning. In this case, a passive sentence is generally followed by a sentential object. These sentences are generally comparable to English sentences such as, "it is believed that smoking is bad for health" or "it is said that apples are good for health." Consider the following sentences:

(30) ɛ̃ kahā jātā hɛ ki nehru bare dayālu ādmī the
said goes aux that nehru very kind man was
It is said that Nehru was a very kind man.

(31) ājkal yeh sunā jātā hɛ ki har deš̆ mẽ̃
now-a-days this heard goes aux that every country in
kapṛo ke dām barh rahe hɛ̃̃
clothes of cost increase prog. aux
Now-a-days it is heard that in every country the prices of clothes are increasing.

Notice that in (30) and (31) the ex-subject is not present and they convey an "impersonal" meaning. The "capabilitative" meaning is not conveyed by these sentences. B. Kachru (1977) suggests that this function of passive sentences (i.e., to express an "impersonal" meaning) has been borrowed from English into Hindi as a result of the language-contact between English and Hindi over a long period of time.

4.0 Volitional act and the Hindi Passive

The discussion in (3.0)--(3.4) shows that passive sentences in Hindi express a variety of meanings. However, whether or not a verb will undergo Passive cannot be judged on the basis of these meanings because none of these meanings is shared in common by all passive sentences in Hindi. Recall that

(i) The "capabilitative" meaning is lost in the subject-less passive sentences (cf. [3.2]).

(ii) The "prescriptive" as well as the "impersonal" meaning is governed by the structure (it has to be a subject-less construction) and tense and aspect (present tense and habitual aspect for the "prescriptive" meaning).

Therefore, neither the "prescriptive" nor the "impersonal" meaning can be said to be the characteristic of the verbs which undergo Passive.

A close examination of passive sentences in Hindi shows that they express a volitional act, in the sense that the subject's volition controls the act expressed by the verb, as has been observed by Kachru (1972) and Hasan (1971). In the following discussion it will be shown that every passive sentence (irrespective of what kind of meaning it conveys) expresses a volitional act. The examples in focus are (3.a), (21) and (30). Notice that (3.a) expresses the "capabilitative" meaning, (21) expresses the "prescriptive" meaning and (30) expresses the "impersonal" meaning. The hypothesis that (3.a), (21) and (30) express a volitional act is supported by the following examples (32)--(34). Examples (32), (33) and (34) are the counterparts of (3.a), (21) and (30) respectively. In (32)--(34) it is stated that the act expressed by the verb took place although the subject did not do it intentionally. In other words, the subject's volition is explicitly negated in those sentences. If our hypothesis is correct, i.e., if the passive sentences express a volitional act, then such negation of the subject's volition must present contradiction. Notice that (32)--(34) do in fact present contradiction.

(32) *laṛke se ciṭṭhiyā̃ paṛhī jātī hɛ̃; jānbūjhkar nahī̃
boy by letters read go aux intentionally not
The letters are read by the boy (but) not intentionally.

(33) *dīwālī par ghar ghar mɛ̃ dīp jalāe jāte hɛ̃; jānbūjhkar nahī̃
dīwālī on every house in lamps lit go aux intentionally not
The lamps are lit on the Dīwālī-day, (but) not intentionally.

(34) *Kahā jātā hɛ --jānbūjhkar nahī̃--ki nehru bare dayālu ādmī the
said go aux intentionally not that nehru very kind man was
It is said (but)--not intentionally--that Mr. Nehru was a very kind man.

Notice that in the first clause in (32)--(34), the subject's volition is negated. Also notice that (32)--(34) are ungrammatical. This ungrammaticality can be explained in a straightforward fashion if we assume that in (3.a), (21) and (30), subject's volition is presupposed; therefore, its negation in (32), (33) and (34) respectively, presents contradiction. The examples (32)--(34) show that the volitional act is characteristic of all passive sentences in Hindi.

4.1 Volitionality and the exception-verbs

The discussion in the following section is focused on two major points: first, whether or not a verb will undergo Passive crucially depends on whether or not it expresses a volitional act. Second, the classification of the verbs into two classes, i.e., verbs expressing a volitional act and verbs expressing a non-volitional act, is independently motivated.

4.1.a. Inanimate subject and the exception-verbs

The verbs whose subject is generally inanimate (and thereby does not have any volition) typically fail to undergo Passive. Notice that verbs

such as khilnā 'to bloom', murjhānā 'to wither', ubalnā 'to boil' (intransitive) and biknā 'to sell', generally take inanimate subjects such as phūl 'flower' (khilnā 'to bloom' and murjhanā 'to wither'), pānī 'water' (ubalnā 'to boil') and kitab 'book' (biknā 'to sell'). Those verbs do not undergo Passive. Consider the following examples:

(35) *phūl se khilā jātā hɛ
flower by bloomed go aux
It is bloomed by the flower.

(36) *Kaliyõ se murjhāyā jātā hɛ
buds by withered go aux
It is withered by the buds.

(37) *pānī se ublā jātā hɛ
water by boiled goes aux
It is boiled by the water.

(38) *ghar se bikā jātā hɛ
house by sold goes aux
It is sold by the house.

Now consider the following verbs which take animate or inanimate subjects. The examples (39) and (40) show that those verbs are passivized only when their subject is animate.

(39) larkī andar ātī hɛ
girl in comes aux
The girl comes in.

(39.a)(larkī se) andar āyā jātā hɛ
(girl by) in came goes aux
It is come in by the girl.

(40) havā andar ātī hɛ
air in comes aux
The air comes in.

(40.a)*havā se andar āyā jātā hɛ
air by in came goes aux
It is come in by the air.

Notice that the verb ānā 'to come' can take either animate (39) or an inanimate subject (40). However, when Passive is applied to (39) and (40) respectively, (39.a) is grammatical, while (40.a) is not. (40) differs from (39) in that the subject of (40) is inanimate (havā 'air'), while the subject of (39) is animate (larkī 'girl'). This discussion shows that in order for the verb to undergo Passive, the subject's volition must control the act expressed by the verb or, in other words, the verb must express a volitional act.

4.2 The negation test

The purpose of this test is to determine, independently of the passive construction, whether or not the subject's volition controls the act expressed by the verb. If the subject's volition controls the act in a sentence, its negation in the conjoined sentence should result in contradiction. Now contrast the verb tornā 'to break' (which undergoes Passive) with the exception-verb girnā 'to fall'. When it is claimed that the action took place against the subject's volition, the use of the regular verb tornā 'to break' produces contradiction, while the use of the exception-verb girnā 'to fall' does not.

tornā 'to break' (regular verb--undergoes Passive)

(41) ??? mɛ̃ ne phuldan tora halaki mɛ̃ tornā nahi cahtī thī
I ag.M. vase broke although I to break not wanted
I broke the vase although I did not want to break it.

girnā 'to fall' (exception-verb)

(42) us bhīṛ mɛ̃ vah larkī gir paṛī halaki vah girnā nahi
that crowd in that girl fell down although she to fall not
cahtī thī
wanted
That girl fell down in that crowd although she did not want to.

Notice that when the volition of the subject mɛ̃ 'I' in breaking the vase is negated in the conjoined sentence, (41) produces contradiction, while (42) is fine when the volition of the subject vah larkī 'that girl' is negated in the conjoined sentence. This result of the negation test shows that the verbs which undergo Passive (i.e., tornā 'to break') express a volitional act while the verbs which fail to undergo Passive (i.e., girnā 'to fall) express a non-volitional act.

5.0 Evidence from constructions involving volitionality

In this section I will present evidence to show that the verbs which fail to undergo Passive, typically fail to participate in the constructions which require the verb to express a volitional act. The regular verbs are contrasted with the exception-verbs and it is shown that the regular verbs can participate in those constructions. However, then the exception-verbs are used, the result is ungrammatical.

5.1 The Imperative Construction

The first one of this type of construction is the imperative construction. It is generally assumed that the subject's volition controls the act expressed by the verb in the imperative construction. Thus in English, in a sentence like 'carry it!', it is assumed that the volition of the implied subject 'you' controls the action of carrying. Also, notice that if a verb does not express a volitional act, it fails to participate in this construction, i.e.,

(43) *Like it!

(44) *Resemble it!

Notice that the subject's volition does not control the act of liking (43) and resembling (44). The following examples show that this construction is blocked for the exception-verbs.

(45) *yād āo! 'be remembered!'

(46) *lago! 'feel!'

(47) *khoo! 'lose yourself!'

Now consider the following examples:

(48) yād karo! 'remember!'

(49) toṛo! 'break it!'

(50) likho! 'write!'

Notice that the verb yād ānā 'to be remembered' (45), lagnā 'to feel' (46) and khonā 'to lose oneself' (47), are exceptions to Passive, while the verbs yad karnā 'to remember' (48), tornā 'to break' (49) and likhnā 'to write' (50) are not exceptions to Passive. This discussion shows that subject's volition does not control the act expressed by the exception-verbs ((45)--(47)), while the subject's volition controls the act expressed by the verbs which undergo Passive ((48)--(50)).

5.2 Koš̃iš̃ karnā 'try to' construction in Hindi

Another construction which requires the verb to express a volitional act is the one in the English sentences such as "Bill tried to help John". Notice that the verbs which participate in this construction are verbs like "to please" (i.e., Jim tried to please his boss), "to kill" (i.e., Sam tried to kill Bill) and not verbs such as "to resemble" (??? John tried to resemble Bill). Here again, the subject's volition is involved in the acts expressed by the verbs of pleasing, killing, and helping while it is not involved in the act expressed by the verbs such as "resemble".

Now let us consider the Hindi verbs of this type. The following examples show that while this construction is blocked for the exception-verbs, such as khonā 'to lose one self', yad ana 'to be remembered', it is not for the verbs which undergo Passive.

(51) sudhā ne sab janne kī kos̃is̃ kī
sudhā ag.M. all to learn tried
Sudhā tried to learn everything.

(52) mɛ̃ ne soneki kos̃is̃ kī
I ag. M. to sleep tried
I tried to sleep.

(53) *dilīp ne khonekī kosīš kī
dilīp ag. M. to lose himself tried
Dilīp tried to lose himself.

(54) *seemā ne yād ānekī kosīš kī
seemā ag. M. to be remembered tried
Seema tried to be remembered.

Notice that the use of verbs such as jānnā 'to learn' (50) and sonā 'to sleep' (52) in this construction, does not produce ungrammaticality. In contrast to this, when used in this construction, the verbs khonā 'to lose one self' (53) and yad anā 'to be remembered' (54), result in ungrammaticality. The verbs jannā 'to learn' (51) and sonā 'to sleep' (52) undergo Passive while khonā 'to lose oneself' (53) and yad anā 'to be remembered' (54) do not. On the basis of this discussion, we can infer that the subject's volition does not control the act expressed by the exception-verbs while it does in the case of the verbs which undergo Passive.

5.3 The dative-subject-verbs and non-volitionality

The verbs which require their subject to be in the dative case are exceptions to Passive in Hindi. A native speaker of Hindi intuitively feels that those verbs do not express a volitional act. Thus for him the subject's volition does not control the act of remembering expressed by the dative-subject-verb yad anā 'to be remembered'. The following evidence supports this intuition. Notice that in (55) and (56), the dative-subject-verb yād anā 'to be remembered' (literally, memory come) is contrasted with the verb yad karnā 'to remember' (literally, to do memory). In (55) and (56) it is stated that the act of remembering expressed by the verbs yād anā 'to be remembered' (to come to memory) (55) and (56) respectively took place against the subject's volition. While (56) presents contradiction, (55) does not.

(55) (mere) na cāhnepar bhī merā dost mujhe bār bār
(my) not wanting in spite of my friend to me again & again
yad atā hɛ
remember aux
Although I do not want to, I remember my friend over and over again.

(56) ??? (mere) na cāhnepar bhī mɛ̃ apne dost ko
(my) not wanting in spite of I my friend Obj, M.
bār bār yād karta hū̃
again & again remember aux
Although I do not want to, I remember my friend over and over again.

The examples (55) and (56) indicate that the subject's volition does not control the act expressed by the dative-subject-verb yād anā 'to be remembered' (55). Therefore, even when the act is said to have taken place against the subject's volition, no contradiction results. In contrast to this, the subject's volition controls the act of remembering expressed by the verb yād karnā 'to remember' (56). Therefore, when the act is said to have taken place against subject's volition, (56) presents contradiction. Notice that the verb yād anā 'to be remembered' is an exception-verb while

yād karnā 'to remember' is not. The hypothesis, i.e., dative-subject-verbs do not express a volitional act, is further supported by the cross-linguistic evidence presented in Krishnamurti (1975), McAlpin (1976), and Sridhar (1976).

5.4 In this section I will present more evidence for the hypothesis that the exception-verbs express a non-volitional act. The Hindi adverbial dhyān se 'carefully' is comparable to the English adverb 'carefully'. Notice that the adverb 'carefully' (in English) can accompany the verbs like "drive" (i.e., He drives carefully) and "listen" (i.e., Listen carefully!), where the subject's volition controls the act of driving and listening respectively. However, notice that "carefully" does not occur with the verbs such as "get hurt" (*John got hurt carefully) or "like" (*I carefully liked the movie), where the subject's volition does not control the act expressed by the verbs "to get hurt" and "to like" respectively. Thus in English, the adverb "carefully" does not accompany a verb which does not express a volitional act. The Hindi data shows that dhyān se 'carefully' accompanies only those verbs which undergo Passive and typically does not accompany the exception-verbs. Consider the following examples:

(57) jān dhyān se partā hɛ
John carefully studies aux
John studies carefully.

(58) *jān dhyān se kho gayā
John carefully lost himself
John carefully lost himself.

Notice that the adverb dhyān se 'carefully' can occur with the verb parhnā 'to study' but not with the verb khonā 'to lose oneself'. While parhnā "to study" undergoes Passive, khonā 'to lose oneself' does not. This shows that the exception-verbs do not express a volitional act whereas the verbs which undergo Passive express a volitional act.

6.0 Conclusion

The major points in the preceding discussion can be summed up as follows: First, that Hindi has a passive rule (section (1.0)). Second, that the list of exceptions to Passive in Hindi consists of both transitive and intransitive verbs. Within the framework of hypothesis I (Lakoff 1965), these verbs will have to be treated as lexical exceptions and be marked as (-Passive) in their lexical entry. Third, passive sentences in Hindi express a variety of meanings under various conditions (sections (3.2)--3.4)). However, all of them share one thing in common, i.e., they all express a volitional act (4.0). Fourth, the discussion in (4.1)--(5.4) reveals that the verbs which undergo Passive express a volitional act while the exception-verbs do not.

On the basis of these observations, the following conclusions can be drawn: The exceptions to Passive in Hindi do not present true irregularity. Passive in Hindi applies to verbs which express a volitional act. Since the exception-verbs typically express a non-volitional act, they fail to undergo the rule and present apparent irregularity. Thus Passive in Hindi

should be labelled as a governed rule in the sense discussed in Green (1974), Green (1976), and Green and Morgan (1976), i.e., it applies to a semantically specifiable class of verbs.

The classification of verbs into two classes (i.e., the verbs expressing a volitional act and the verbs expressing a non-volitional act) is independently motivated (recall the discussion in sections (5.0)--(5.4). The verbs belonging to these classes behave differently from each other with regard to other syntactic constructions in Hindi. Thus this classification enables us to explain the "exceptional" behavior of the exception-verbs and captures the generalization, namely, that the exception-verbs are exceptions to the constructions which require the verb to express a volitional act. The hypothesis which marks these exceptions as true irregularity, misses this generalization.

The hypothesis proposed in this paper explains why some borrowed verbs (i.e., verbs borrowed from English) undergo Passive while some others do not. Notice, Hindi has borrowed several nouns from English, e.g., ɛkting (acting), ʂ̃ak (shock), etc. (For further discussion see B. Kachru: 1977). Some "hybrid" compound-verbs are formed with these nouns. The process of deriving those compound-verbs is as follows: An operator (a native Hindi verb) is attached to the word borrowed from English; these two (the borrowed noun and the operator) function as a unit--a compound-verb. Thus the following verbs are formed: ɛkting karnā 'to act' or 'to pretend'. Notice that in this compound-verb, ɛkting 'acting' is the borrowed noun and karnā 'to do' is the Hindi verb which is used as an operator. Now consider the other compound-verb ʂ̃ak lagnā 'to get a shock' or 'to be shocked', in which ʂ̃ak 'shock' is the borrowed noun and lagnā 'to feel, to be attached' is the Hindi verb used as an operator. While ɛkting karna 'to act' is passivizable, ʂ̃ak lagnā 'to be shocked' is an exception to Passive. Under hypothesis I, the grammar of Hindi will have to mark ʂ̃ak lagnā 'to be shocked' as an exception to Passive. However, it will not explain why some of the verbs formed with the borrowed nouns undergo Passive while some do not. In the framework of our hypothesis, this situation can be explained in a straightforward fashion. Notice that the exception-verb ʂ̃ak lagnā 'to be shocked' is a dative-subject-verb. The discussion in (5.3) shows that the dative-subject-verbs express a non-volitional act and therefore it is only expected that it will not undergo Passive. Also, the behavior of these verbs with regard to the constructions discussed in (5.1), (5.2) and (5.4) shows that while ɛkting karnā 'to act' expresses a volitional act, ʂ̃ak lagnā 'to be shocked' does not.[3]

6.1 The discussion in this paper raises some theoretical questions:

It has been observed (Green and Morgan 1976) that all rules that change grammatical relations are governed. However, it is not clear as to what degree of relation-changing capacity is necessary for a rule to be governed. The Hindi Passive changes the grammatical relations only partially. Notice that the ex-object in the Passive sentence (9) controls reflexivization and thereby provides evidence for object-promotion. However, (9) does not provide evidence for subject-demotion (i.e., the ex-subject continues to control reflexivization in (9)). Kachru et al. (1976) have shown that the subject of the passive sentence in Hindi is less subject-like. On the whole, the Hindi Passive only partially changes the grammatical relations.

The discussion in this paper shows that Passive is a governed rule in Hindi. Can we then say that the rules which only partially affect the grammatical relations are also governed? This assumption will give rise to the following questions:

(a) Are grammatical relations always discrete?

(b) What degree of relation-changing capacity of a rule is necessary in order to establish a correlation between relation-changing rules and rule-government?

The "prescriptive" meaning and the "impersonal" meaning conveyed by some passive sentences in Hindi provide evidence for the role of language-use in determining the meaning of a sentence. This raises the question, how did Hindi acquire the "capabilitative" meaning? Is it a function of language-use or of the linguistic structure?

The Hindi Passive shows that the syntactic behavior of the verbs is determined by their role in the semantic structure of the language. One might therefore ask, to what extent semantic structure controls the syntactic structures of a language? At present these are open questions and further investigation is needed.

NOTES

[1]I am greatly indebted to Professors Yamuna Kachru, Braj Kachru and Jerry Morgan for their invaluable suggestions, criticisms and guidance but for which this paper would be much worse than it is. I claim the responsibility for the shortcomings that remain.

[2]The influence of English is seen at various linguistic levels (i.e., lexicon, morphology and syntax) in the newspaper register of Hindi. "A number of syntactic patterns which are transferred from English have a high frequency in the newspaper register such as in sports news and in headings and captions" (B. Kachru 1977). Kachru points out two major reasons for the direct influence on this particular register of Hindi "one reason being that until recently, English was used as the main source for translating news items into Hindi, since the facilities for direct communication of news in Hindi (say, teleprinters in Hindi) were not available. Second, a large number of successful English newspapers started sister publication in Hindi (i.e., The Hindustan Times and Hindustan (New Delhi) etc.)."

In Hindi, traditionally, active forms of verbs are used to convey the "impersonal" meaning. However, the following paragraphs from newspapers illustrate the use of passive to convey the "impersonal" meaning in the newspaper register:

(a) yah bhī mãg kī gaī ki ukt samsthān mẽ dhan kī
this also demand did went that said province in finances of
herāpherī kī jac kī jāy
mishandling of investigation did go (optat.)
It has been demanded that there be an investigation of the mishandling of finances in the said province. (Navbharat Times, New Delhi, October 19, 1977).

(b) yahã̃ tak kahā gayā ki līnā ke ṭhīk hote hī use aspatāl
here upto said went that line of recovering soon her hospital
se sidhe kahī̃ ɛsī jagah bhej diyā gayā . . .
from straight some such place sent went
It was even said that soon after her recovery Lina was sent from the hospital to some place . . . (Saptahik Hindustan, May 22, 1977).

[3]There is evidence to show that the verb ɛkting karnā 'to act' expresses a volitional act; in contrast to this, šak lagna 'to get a shock' does not express a volitional act. Notice that ɛkting karnā 'to act' participates in the constructions which require the verb to express a volitional act, while šāk lagnā 'to get a shock' fails to participate in those constructions. Consider the following examples:

(a) The Imperative construction (cf. 5.1)
1. ɛkting karo 'act!, do the acting!'
2. *šāk lago 'get a shock!'

(b) kośiś karnā 'try to' (cf. 5.2)
3. rām ne ɛkting karne kī kośiś kī
Rām tried to do the acting.
4.*rām ne šak lagne kī kośiś kī
rām ag.M. shock get tried
Rām tried to get a shock.

(c) dhyān se 'carefully' (cf. 5.4)
5. vah dhyān se ɛkting karta hɛ
he carefully acting does aux
He does the acting carefully.
6.* vah dhyān se šāk lagtā hɛ
usko
he/him gets a shock carefully.

Notice that constructions such as Imperative, 'Try to' and 'carefully' are blocked for the verb šāk lagnā 'to get a shock' (recall (2, 4, and 6)). On the other hand, those constructions not blocked for ɛkting karnā 'to act' (1, 3, and 5). On the basis of this, we can conclude that while ɛkting karnā 'to act' expresses a volitional act, the verb šāk lagnā 'to get a shock' does not. Also, the fact that šāk lagnā 'to get a shock' is a dative-subject-verb indicates that it expresses a non-volitional act. Our hypothesis correctly predicts that while ɛkting karnā 'to act' will undergo Passive, šak lagnā 'to get a shock' will not.

REFERENCES

GONDA, J. 1975. Vedic literature. Otto Harrassowitz. Wiesbaden.

GREEN, G. 1974. Semantics and syntactic regularity. Indiana University Press.

_____. 1976. The notion of rule government. Problems in linguistic matatheory. Department of Linguistics. Michigan State University.

_____ and MORGAN, J. 1976. Notes toward an understanding of Rule Government. Studies in Linguistic Sciences. Urbana; University of Illinois. Volume 6, Number 1.

GURU, K. P. 1920, 1952. Hindi Vyakran. Kashi Nagri Pracarini Sabha. Varanasi.

HALLIDAY, M. A. K., MCINTOSH, A., and STREVENS, P. 1964. The Linguistic sciences and language teaching. Indiana University Press. Bloomington.

HASAN, R. 1971. Syntax and semantics: Biological and Social factors in Psycho-linguistics. Ed. John Morton, London.

KACHRU, B. 1977. Englishization of Hindi. Linguistic Method. Papers in honor of Herbert Penzel. Ed. I. Rauch et al. De. Ridder Press. The Hague.

KACHRU, Y. 1966. An introduction to Hindi syntax. Urbana; University of Illinois, mimeographed.

_____. 1972. Notes on grammatical categories and participant roles in Hindi-Urdu (unpublished).

_____, KACHRU, B., and BHATIA, T. 1976. The notion "subject": discussion of three North Indian languages. The notion of Subject in South Asian Languages. Ed. M. K. Verma. University of Wisconsin Publication No. 2.

KELLOGG, S. H. 1955. A grammar of the Hindi language. Rouledge and Kegan Ltd. London.

KEITH, A. B. 1928. A history of Sanskrit literature. Oxford.

KRISHNAMURTI, Bh. 1975. Verbs of cognition in Telugu. Osmania Papers in Linguistics, 1:1. Hyderabad. Osmania University.

LAKOFF, G. 1965. On the nature of syntactic irregularity. Indiana University dissertation. Published 1970 as Irregularity in Syntax. Holt.

LAKOFF, R. 1968. Abstract syntax and Latin complementation. M.I.T. Press.

MCALPIN, D. 1975. Dative subjects in Malayalam. The notion of subject in South Asian languages. Ed. M. K. Verma. University of Wisconsin. Madison publication series, publication 2.

MCGREGOR, R. S. 1972. Outline of Hindi Grammar. Clarendon Press. London.

PANDHARIPANDE, R. 1976. On the syntax and semantics of the yā-passive (present) construction in Śatapatha Brāhmaṇa (unpublished).

POSTAL, P. and PERLMUTTER, D. 1974. Relational grammar (unpublished).

Shatapatha Brāhmaṇa. 1967. Volumes I, II and III. Published by The Research Institute of Ancient Scientific studies, New Delhi.

SRIDHAR, S. N. 1976. Dative subject, rule government, and relational grammar. Studies in Linguistic Sciences, Vol. 6, No. 1.

VAJPEYI, K. D. 1966. Hindi Shabdanuahasan. Kashi Nagri Pracarini Sabha. Varanasi.

Studies in the Linguistic Sciences
Volume 8, Number 1, Spring 1978

LANGUAGE-LIKENESS

Irmengard Rauch

Semiotics, the science of signs, derives its recent linguistic heritage from standard bearers such as Saussure, Sapir, Bloomfield, Jakobson and, indeed, Chomsky. Bloomfield's (1939) claim that 'Linguistics is the chief contributor to semiotic' and Jakobson's (1971) hypothesis that '...any human communication of non-verbal messages presupposes a circuit of verbal messages' indicate a modus operandi, namely the function in which all of semiotics displays an inlay of language, in which non-verbal communication is LANGUAGE-LIKE. It is proposed that the isolation and identification of discrete elements shared by idiomorphic semiotic systems and language semiotic systems provide the cardinal data. The universal primitive of LENGTH is extracted in both a set of language examples and a set of music/architecture examples. The LANGUAGE-LIKENESS of the music and architecture semioses is found to reside in the processing of the primitives via the inferential function of abduction -- a function peculiar to language.

Linguistic theory, which is currently in the midst of a new wave of iconoclasm, appears on the one hand to be leaning toward the laboratory sciences for insights into the composition and linguistic functions of the brain, while on the other hand the probings of linguists resemble ever more the speculations of philosophy. The duality is understandable, since language is both physical and mental. Several schools of philosophy claim language as their principal subject matter, thus providing a ready tap for linguists whose sole object is, in fact, language. More immediately, within our own recent linguistic heritage we can uncover an appropriate philosophical framework suggested by none other than Saussure, continued by such standard bearers as Sapir, Bloomfield, and Jakobson, and even endorsed by Chomsky. This, of course, is the discipline of semiotics, the science of signs.

Saussure proposed in his Course in general linguistics (1966:16) 'A science that studies the life of signs within society...' He says

> 'I shall call it semiology. Semiology would show what constitutes signs, what laws govern them. Linguistics is only a part of the general science of semiology; the laws discovered by semiology will be applicable to linguistics, and the latter will circumscribe a well-defined area within the mass of anthropological facts.'

Sapir (1958:104) saw that '...every cultural pattern and every single act of social behavior involves human communication in an explicit or implicit sense.' The anthropological theorist Claude Lévi-Strauss formulated very

clearly a concept of society with an integrated theory of communication consisting of three types of exchange: exchange of messages (linguistics), exchange of goods (economics), and exchange of mates (kinship). Thus semiotics is, in effect, a science studying communication within society. We will assume, for the moment, that this is society within the species homo sapiens and, for the major part of this paper delay reference to semiotics outside this species, i.e. zoosemiotics. Nevertheless, within human semiosis, we must account for all three types of Lévi-Strauss' communication, which, in turn, have many subtypes.

Let us return to our linguistic heritage. Bloomfield tells us in his Linguistic aspects of science (1939:55) that 'Linguistics is the chief contributor to semiotic.' How is this possible if we are to understand, according to Sebeok (1975:95), that 'Semiosis, independent of form or substance, is...seen as a universal, criterial property of animate existence'? In other words, what do we do with that communication or semiosis which is not the chief contributor to semiotics? Jakobson (1971:662) writes '...the subject matter of semiotic is the communication of any message whatever, whereas the field of linguistics is confined to the communication of verbal messages.' Temporarily he answers our question: Linguistics '...has a narrower scope, yet, on the other hand, any human communication of non-verbal messages presupposes a circuit of verbal messages, without a reverse implication.' Thus Lévi-Strauss (1963:83) claims for his three types of communication or exchange that '...language comes into play at all levels', i.e. it preexists each exchange, it accompanies each, and it serves as a means of translation for each non-verbalized exchange. This last function serves as the focal point of the present paper: We want to know how language serves as a means of translation for non-verbalized exchange, how non-verbal messages presuppose a circuit of verbal messages, i.e. how all of semiotics has an INLAY of language, how non-verbal communication is LANGUAGE-LIKE.

The centrality of linguistics prompted Roland Barthes (1970:11) to reverse Saussure's statement ('Linguistics is only a part of the general science of semiology' cf. above) by the statement '...linguistics is not a part of the general science of signs, even a privileged part; it is semiology which is a part of linguistics....' However, upon close inspection we see that Barthes overstated his case, and that, in fact, he distinguished language which is the object of linguistics from what he termed 'second-order language,' which is the object of semiotics.

Clearly semiotics and linguistics are independent sciences, but just as clearly they are intimately related, since language provides an inlay for all modalities of semiotics. When we observe human language in action, we observe an organism (a human being) producing an expression (a sound, orthographic sequence, gesture) in order to refer by it to something (an object) (Carnap 1961:8-9). Thus the language act is comprised of (1) a speaker, (2) an expression, and (3) a designatum, which correlate directly with the tripartite division of semiotics, viz. pragmatics, syntactics, and semantics, respectively. For some linguists the semiosis of verbal language appears to be a challenge easily commensurate with life-time devotion, so that they opt not to treat data from speech surrogates such as cryptography, whistle speech, drumbeat communication, and so-called paralinguistic structures such as laughing, crying, whispering, and noises of the tongue and

lips associated with interjections. Still more linguists exclude from their domain idiomorphic semiotic data, i.e. data from semiotic systems which are not variform direct superstructures on languages. Samples of idiomorphic systems can be readily named by reference to the senses, thus kinesics and proxemics relate to the senses of touch and sight, music to the auditory sense, architecture to sight, odor to the olfactory sense, culinary semiotics to the sense of taste as well, and so forth (cf. further Rauch forthcoming a).

Let us consider the likeness between e.g. music and language, or architecture and language? Saussure (1966:17) simply says

> '...the characteristic that distinguishes semiological systems from all other institutions shows up clearly only in language where it manifests itself in the things which are studied least,...'

That was around 1915, and apparently we are still groping to identify those least studied things, to which Saussure ascribes the semiotic mechanism, but never mentions as such. Somewhat parallel is the proposal of Jakobson (1971:703) that

> 'The cardinal functions of language--referential, emotive, conative, phatic, poetic, metalingual--and their different hierarchy in the diverse types of messages have been outlined and repeatedly discussed. This pragmatic approach to language must lead mutatis mutandis to an analogous study of the other semiotic systems....'

But we ask, what is this mutatis mutandis? It is precisely the crux of the problem. One is then not surprised to read as recently as 1975 in an article entitled 'Spoken and unspoken meanings' by Cohen (24), that 'It remains quite an open question whether this clearly recognizable domain of study [the semiotic domain] will lend itself to the construction of non-trivial, empirically testable theories at some appropriately deep level of explanatory generality' (cf. further Rauch 1977).

It seems that the comparison among diverse modalities of massive or large scale relationships, as represented perhaps in the cardinal functions of language (above) is somewhat premature. The linguist may have more success in relying on what is recognized as a principal time-proven tool of all modern linguistic methodologies, viz. their ability to isolate and identify discrete elements. Of the several theories of semantic development, the universal primitives hypothesis, recently proposed especially by Postal (1966:fn. 10) and extended by Bierwisch (1967:1-4, 34-6 and 1970: 181-2), appears in particular to link with semiotic theory. In effect, Postal and Bierwisch hold that the real world, or input into the human organism, is interpreted in the form of universal semantic components or primitives in a species specific manner.

Let us accordingly choose a semantic primitive and identify it in several modalities, in order to understand how it is interpreted into language in non-linguistic semiotic systems. The primitive which we choose is

LENGTH or DURATION, pinpointed by Miller and Johnson-Laird (1976:27) as a universal. To be sure it is readily identifiable in linguistic semiotics, but even within linguistic data, LENGTH is more transparent in certain structures than in others. Thus, in the semantic component, e.g. in the sentence 'A mile is longer than a kilometer,' the perception of LENGTH is more direct than on the syntactic level, which may be exemplified by gapping in e.g. 'John reads Dickens, Joe Hemingway, and Jill Kant' as opposed to 'John reads Dickens, Joe reads Hemingway, and Jill reads Kant.' A morphological example of LENGTH often cited in semiotic literature for its iconicity is the non-suppletive comparison of adjectives in Indo-European, where the positive, comparative, and superlative suffixes mirror a stratified phoneme increment, thus e.g. old, older, oldest. On the phonological level LENGTH is, of course, perceptible where it alone distinguishes one word from another; thus in Kamba, a Bantu language, three degrees of vowel LENGTH are contrastive, e.g. kwelela 'measuring', kwele·la 'moving backwards and forwards', kɔɔle:la 'aiming at.' Consonantal LENGTH is observed in Italian fato 'fate' versus fat:o 'done.' We are, of course, familiar with expressive or affective LENGTH as in the English expression too bad. This leads us technically into what have been considered borderline linguistic levels, specifically what Trager (1958) has called 'paralanguage.' For example, his so-called voice qualifier of extent, which would oppose 'drawl' to 'clipping', represents the feature LENGTH on the paralinguistic level.

Leaving language and paralanguage but still dealing with acoustic LENGTH, let us cross over to the semiotic system of music. An example is the choral sequence of the first twelve notes of equal LENGTH in the Ode to Joy of Beethoven's Ninth Symphony. Excepting performance modifications, any one of the twelve notes which might be played shorter or longer relative to its adjacent notes would be perceived as incorrect. Not surprisingly, Gerald Warfield, who in his Layer analysis (1976) explicates the highly structured principles underlying tonal music, stresses that (xi) '...music can be (and is, most of the time) understood intuitively....'

Before considering how musical LENGTH is language-like, let us apply the LENGTH primitive to a very concrete semiotic system, viz. architecture, and specifically to the basic box, which is man's most common form of construction, be it a house, a skyscraper, or whatever. Miller and Johnson-Laird (1976) distinguish between two systems in conceptualizing an object in space, 1) a deictic system which is relative to the speaker's ego location and the coordinate axes of an object, and 2) an intrinsic system which is relative to the coordinate axes of the object itself. In the case of the box, if the horizontal dimension, labeled A, is its maximal dimension, it is intrinsically the 'length' or 'longer than' the secondary horizontal dimension, labeled B, which is its 'width' or 'shorter than' A. Interestingly, Miller and Johnson-Laird wrote (404): 'Recognizing the intrinsic nature of the parts of some objects seems to be something that users of English manage to do without self-consciously conceptualizing it.' Once more we superimpose the question: How then, is the LENGTH feature in architecture language-like?

It is possible that structure or component analogues between linguistic semiotic systems and non-linguistic semiotic systems may prove insightful, e.g. Uspenskij's (1973) level of articulation of an icon as an analogue to

the phonological level in natural languages. This is a level of figuration which corresponds to the most general painting processes, which are in themselves non-significant, but which allow the painter to represent meaningful complexes.

However, beyond such analogues the primitive LENGTH, whether in the visual arts, music, paralanguage, or language, is ultimately interpreted as a perceptual judgment which derives solely from that semiotic system alone capable of propositions, i.e. language. Postal (1966:fn. 10) argued that '...the relation between the semantic primitives and their combinations which are part of the combinatorial structure of language and the world is not learned but innate.' Bierwisch (1970:181) hypothesized that '...all semantic structures might finally be reduced to components representing the basic dispositions of the cognitive and perceptual structure of the human organism.'

Now this innate ability of the mind to process semantic primitives can find an explanation in the theory of inference of Charles Sanders Peirce, the co-founder with Saussure of modern semiotic. According to Peirce we experience external objects as percepts and interpret them through a cognitive operation. All cognition involves inference and the type of inference which Peirce holds to be the most common kind of reasoning is abduction, whereby the minor premise is derived from the major premise and the conclusion. Our first premises are perceptual judgments, viz. the formation of a mental proposition concerned with our sense experience when making the judgment. Peirce defines this innate manner of interpreting an experience or a perceived object as '...a judgment asserting in propositional form what a character of a percept, directly present to the mind, is' (1960 V:par. 54). Meaning thus originates in the perceptual judgment, which itself can be regarded as an extreme case of abductive inference. The term 'educated guess' closely approximates the abductive process, which Peirce claims is instinctive. The interpretation of our chosen primitive LENGTH then, as displayed in the various semiotic modalities, derives from one and the same innate ability of judgment and hypothesis formation, an ability which belongs only to the semiotic modality of language.

Thus the LENGTH cognition of the non-linguistic modalities presented in this paper is language-like in that it proceeds from reasoning of the following sort:

I. for the musical semiosis:

Each of the twelve notes equals one beat.
E is one of the twelve notes.
E equals one beat.

II. for the architectural semiosis:

Maximal dimension equals the length.
A is the maximal dimension.
A equals the length.

Observe that the middle statements are abductively inferred.

Before concluding let us turn briefly to zoosemiotics, purposely excluded thus far. Peirce (1960 VI:par. 491) compared man's instinctive power of so-called guessing with the instincts of animals for feeding and breeding. Sebeok (1972:46-7) describes what we may discern as differences of LENGTH of the bee dance among six races of bees. The hundred meter distance between the hive and the food source is calculated at six different LENGTHS in revolutions per fifteen seconds: Italian 7.95, Austrian 8.4, German 9, Punic 9.05, Caucasian 9.8, Egyptian 9.25. Man's observation of the LENGTH differential is, of course, not the issue, and Sebeok rightly reports that an encounter between e.g. an Austrian bee and an Italian bee would precipitate confusion of signals.

Purposely excluded also has been any species designation for the word language, i.e. we have refrained from writing human language. Accordingly, the questions which arise directly from the present paper: Can one even contemplate a language-inlay in non-verbal semiosis? and if so, What is the language-likeness in the animal code? join those provocative questions presently stimulating the necessary and inevitable reappraisal of the concept of language itself (cf. Rauch forthcoming b).

To conclude, it appears appropriate to turn to Chomsky's little known endorsement of the Peircean theory of inference referred to at the beginning of this paper. In Language and mind (1972:93) Chomsky writes:

> "Speculating about the future, I think it is not unlikely that the dogmatic character of the general empiricist framework and its inadequacy to human and animal intelligence will gradually become more evident as specific realizations, such as taxonomic linguistics, behaviorist learning theory, and the perception models, heuristic methods, and 'general problem solvers' of the early enthusiasts of 'artificial intelligence' are successively rejected on empirical grounds, when they are made precise and on grounds of vacuity when they are left vague. And--assuming this projection to be accurate--it will then be possible to undertake a general study of the limits and capacities of human intelligence, to develop a Peircean logic of abduction."

REFERENCES

BARTHES, Roland. 1970. Elements of semiology. Translated by A. Lavers and C. Smith. Boston: Beacon.

BIERWISCH, Manfred. 1967. Some semantic universals of German adjectivals. Foundations of Language 3.1-36.

_____. 1970. Semantics. In J. Lyons, ed.: New Horizons in linguistics, pp. 166-84. Harmondsworth, Middlesex: Penguin Books.

BLOOMFIELD, Leonard. 1939. Linguistic aspects of science. (International encyclopedia of unified science, vol. 1, no. 4.) Chicago: University of Chicago Press.

CARNAP, Rudolf. 1961. Introduction to semantics. Cambridge, MA: Harvard University Press.

CHOMSKY, Noam. 1972. Language and mind. New York: Harcourt, Brace, Jovanovich.

COHEN, J. Jonathan. 1975. Spoken and unspoken meanings. In T. A. Sebeok, ed.: The tell-tale sign: a survey of semiotics, pp. 19-25. Lisse: Peter de Ridder Press.

JAKOBSON, Roman. 1971. Roman Jakobson: selected writings II. The Hague: Mouton.

LEVI-STRAUSS, Claude. 1963. Structural anthropology. Translated by C. Jacobson and B. G. Schoepf. New York: Basic Books.

MILLER, George A. and JOHNSON-LAIRD, Philip N. 1976. Language and perception. Cambridge, MA: Belknap Press of Harvard University Press.

PEIRCE, Charles S. 1960. Collected papers of Charles Sanders Peirce. vol. V and VI. Cambridge, MA: Belknap Press of Harvard University Press.

POSTAL, Paul M. 1966. Review article of Elements of general linguistics, by André Martinet. Foundations of Language 2.151-86.

RAUCH, Irmengard. 1977. Review of The tell-tale sign: a survey of semiotics, ed. by T. A. Sebeok. Lingua 41.183-89.

_____. forthcoming a. Semiotics and language. Language Quarterly (University of South Florida).

_____. forthcoming b. The language-inlay in semiotic modalities. Semiotica.

SAPIR, Edward. 1958. Selected writings of Edward Sapir in language, culture and personality, ed. by D. G. Mandelbaum. Berkeley: University of California Press.

SAUSSURE, Ferdinand de. 1966. Course in general linguistics, ed. by C. Bally et al. Translated by W. Baskin. New York: McGraw-Hill.

SEBEOK, Thomas A. 1972. Perspectives in zoosemiotics. (Janua linguarum, series minor, 122.) The Hague: Mouton.

_____. 1975. Zoosemiotics: at the intersection of nature and culture. In T. A. Sebeok, ed.: The tell-tale sign: a survey of semiotics, pp. 85-95. Lisse: Peter de Ridder Press.

TRAGER, George L. 1958. Paralanguage: a first approximation. Studies in Linguistics 13.1-12.

USPENSKIJ, Boris A. 1973. Per l'analisa semiotica delle antiche icone russe. Ricerche semiotiche: nuove tendenze della scienze umane nell' URSS, ed. by J. Lotman and B. A. Uspenskij, Italian edition ed. by C. S. Janovič. Translated by C. S. Janovič et al. Turin: Einaudi.

WARFIELD, Gerald. 1976. Layer analysis: a primer of elementary tonal structures. New York: David McKay.

Studies in the Linguistic Sciences
Volume 8, Number 1, Spring 1978

A REANALYSIS OF THE CLASS 5 PREFIX IN SHONA[1]

Pamela S. Silver
Scott R. Krause

This paper is intended to refute the traditional ways of accounting for the word-initial voicing alternations in the class 5/6 noun-adjective series of Shona. We will show that none of the traditional analyses is correct and that the voiced/voiceless phenomenon under discussion should be looked upon as the result of a morphophonemic rule. In other words, we are proposing that there is neither a segmental nor a Ø prefix in class 5, and that our morphophonemic analysis provides a more concrete solution.

When examining any corpus of Shona data, one is immediately struck by the occurrence of a voiceless/voiced alternation between pairs of consonants in certain paradigms, e.g.,

(1)	class 5	class 6	
	ɓadza[2]	mapadza	'hoe'
	bveni[3]	mapfeni	'baboon'
	ɗeŋga	mateŋga	'sky'
	gumbo	makumbo	'leg'
	ǰira	mačira	'blanket'
	bveni ɗoko	mapfeni madoko	'small baboon'
	ǰira ǰena	mačira mačena	'white blanket'
	daka gukutu	madaka makukutu	'hard mud'
	guɗo ɗema	makuɗo matema	'black baboon'
	ɗare dzvuku	matare matsvuku	'red ball'

Traditionally, this voicing alternation has been looked upon as the result of either:

(2) a) a prefix being lost in the NOUN/ADJECTIVE series of class 5;
b) a morpheme (ɣ, ɬ, etc.) which has no overt phonological shape in class 5 but which shows itself in the voicing of certain consonants (Fortune, 173; Fivaz and Ratzlaff 72, etc.);
c) the "prefix" for the NOUN/ADJECTIVE series of class 5 actually being voicing (Heny, 211), or some similar claim.

We hope to show in the following paper that none of these analyses provides a correct synchronic explanation and that the voiced/voiceless phenomenon under discussion be looked upon as the result of the following morphonemic rule:

(3) $$\begin{bmatrix} C \\ +\text{stop} \\ -\text{vc} \\ +5 \\ +\text{native} \end{bmatrix} > [+\text{vc}] / \#\_\_\_\_\_$$

Our proposal, then, is that there is no prefix for the class 5 noun/adjective series.

We shall proceed as follows: First we will examine the distinctive features and environment of the proposed rule and show why each one is necessary. Second, in the course of the discussion, we will examine a few features which might be proposed for our rule and show why these are unnecessary or erroneous in nature.

As we add more and more forms to our list of data, we note that the observed voicing alternation does not occur with every possible voiced/voiceless pair of consonants that exist in Shona. That is to say, the voicing alternation only occurs with phonemes which are [+stop]:

(4)

b/p	bv/pf
d/t	dz/ts
g/k	dǯ/tš (ǰ/č in our notation)

class 5	class 6	
ɓadza	mapadza	'hoe'
ɗare	matare	'bell'
guyo	makuyo	'wooden mortar'
bvupa	mapfupa	'bone'
dzimba	matsimba	'footprint'
ǰira	mačira	'blanket'

But such an alternation does not occur with fricatives:

(5)

v/f	z/s	ž/š

class 5	class 6	
šawi	mašawi	'possessing spirit'
šoko	mašoko	'word'
simba	masimba	'strength'
sondo	masondo	'week'
viri	maviri	'wheel'
vudzi	mavudzi	'hair'
zai	mazai	'egg'
zita	mazita	'name'
fuza	mafuza	'idiot'
fatčke	mafatčke	'cask'

Thus it appears that since non-stops do not show the voicing alternation, and since all stops or stop-initial phonemes do, the most natural way to account for the facts at this point in our discussion is to set up a voicing alternation between stops.

(6)

$$\begin{bmatrix} C \\ +\text{stop} \\ -\text{vc} \end{bmatrix} \sim \begin{bmatrix} C \\ +\text{stop} \\ +\text{vc} \end{bmatrix}$$

Crucial evidence is lacking as to whether affricates should be analyzed as a single [+stop] phoneme (Chomsky and Halle, 317), or as consonant clusters consisting of two separate phonemes. If the second alternative

were chosen, a rule of voicing assimilation would be necessary, i.e.,

$$(7)\quad \begin{matrix} C \\ [\alpha vc] \end{matrix} > [\beta vc] / C \_\_\_\_\_ \begin{matrix} \\ [\beta vc] \end{matrix}$$

Since the evidence needed to determine the status of affricates is lacking, we will adopt Chomsky and Halle's framework, and will refer to affricates as $\begin{bmatrix} +\text{stop} \\ +\text{delayed release} \end{bmatrix}$.

If [+voicing] is the morphological marker for class 5, as proposed by Heny, Fivaz and Ratzlaff, etc., why is it that only stops are affected? In all other classes, all words, regardless of their initial consonants (or indeed vowel), receive the prefix associated with that class. For example, all class 6 nouns and adjectives are prefixed by #ma-. Because of this we would expect all initial consonants in class 5 to become voiced, if in fact the class 5 prefix were [+voicing] or some similar abstraction. That this is not the case, as has been seen from the preceding data, argues against such proposals.

So far we have given evidence that there is a voicing alternation between stops in classes 5 and 6. The next logical question is in which direction this change takes place, i.e. [+vc]>[-vc], or [-vc]>[+vc]. The evidence seems to point toward [-vc]>[+vc] for several reasons. There are cases of

(8) a) class 5 [+vc] ~ class 6 [+vc]

alternations, as in the following,

class 5	class 6	
ɓasa	mabasa	'work'
ɗutu	madutu	'strong wind'
gumbeze	magumbeze	'blanket'

as well as

b) class 5 [+vc] ~ class 6 [-vc]

alternations, as in the following.

class 5	class 6	
ʄadza	mapadza	'hoe'
ɗura	matura	'granary'
gomo	makomo	'hill'

Outside of the affricates and borrowed words, both of which we will discuss below, there are no cases of

c) class 5 [-vc] ~ class 6 [-vc]

alternation, and there are no cases of

d) $\begin{matrix}\text{class 5}\\ [-\text{vc}]\end{matrix} \sim \begin{matrix}\text{class 6}\\ [+\text{vc}]\end{matrix}$

alternations in the language whatsoever.

As such, the voicing distinction between stops has been virtually neutralized in root-initial position in class 5. To take the tack $[+\text{vc}] > [-\text{vc}]$ in class 6 would necessitate an explanation not only of why there are no voiceless stop-initial class 5 words, but also an explanation of exactly in what environment a voiced stop devoices, which is impossible to state even when taking tonology into consideration (i.e., would require marking each item which undergoes this devoicing in the lexicon).

Some linguists may claim, however, that admission into class 5 is dependent upon a root-initial voiced consonant. This assumption is inherently claiming that there is a devoicing rule for class 6 (exactly the opposite of our proposal - a voicing rule for class 5). In view of the fact that a devoicing environment for class 6 is impossible to state, a linguist making such a claim would have to say that all root-inital stops that do not devoice in class 6 are exceptions. However, in view of the fact that an initial voiced consonant is not a valid requirement for inclusion into class 5, since root-initial voiceless fricatives occur in abundance[4], and in view of the fact that our analysis (whereby a voiceless stop voices in class 5) is exceptionless in that all stops voice in the following environment,

(9) /#_____

the $[-\text{vc}] > [+\text{vc}]$ directionality does indeed seem preferable. (This voicing alternation only occurs in root-initial position, as we have seen in the data thus far and will see in greater detail later.) Thus, in the view of simplicity and the lack of convincing evidence to the contrary, we tentatively have

(10) $\begin{bmatrix}\text{C}\\ +\text{stop}\\ -\text{vc}\end{bmatrix} > [+\text{vc}] / \#\_\_\_\_\_$

We have implicitly assumed up to this point in the discussion that the voiced/voiceless alternation only occurs in classes 5 and 6. This is not entirely true, but as we shall see below, all those voicing alternations outside of classes 5 and 6 may be accounted for by a separate rule, 5 and 6 being the only place where our proposed non-phonemic alternation takes place.

The next question is how many morphological features we need to include in our rule, i.e., $[+5]$, $\begin{Bmatrix}+\text{N}\\ +\text{Adj}\end{Bmatrix}$, $[-\text{V}]$, etc. Since this particular morphophonemic alternaion only takes place in class 5 nouns and adjectives, it seems that in the strictest sense both the feature $[+5]$ and $\begin{Bmatrix}+\text{N}\\ +\text{Adj}\end{Bmatrix}$

might not be needed. In the 21 noun classes of Shona, all 21 agreement series have overt prefixes, and in the noun/adjective series, all classes except 5 and 9 have overt prefixes, although it appears that some sort of synchronic nasal may be posited for class 9.[5] As such, the only place in which we have word-initial non-prefix consonants of either series of noun classes is the class 5 series of nouns and adjectives. Therefore, to state both [+5] **and** [{+N / +Adj}] would be redundant, but to state one or the other is necessary, because outside of the noun class series, this rule is not productive. The following data will show that the root-initial stop of verbs, even when in word-initial position (this only happens in the imperative), never voices.

(11)

infinitive	imperative	
kuteŋga	teŋga	'buy'
kutarisa	tarisa	'look'
kupa	ipa	'give'
kukura	kura	'grow'
kučeŋgeta	čeŋgeta	'save'
kupfugama	pfugama	'kneel'

The voicing alternation does seem to serve a function of sorts in class 5, although it is not as pervasive as one might hope if it were a purely morphological function, as other scholars would have it. So we now have

(12) $C > [+vc] / \#\_\_\_\_\_$ where C = [+stop, -vc, +5]

Thus far, our voicing rule has been formulated so as to affect only word-inital voiceless stops in class 5 nouns and adjectives, since we have already seen that verbs are not affected by our rule. It must now be noted that there are two different types of adjectives, and the following may shed more light on the subject. (In the predicate adjectives, as we have elected to call them, #r- in class 6 is an agreement marker; -a- is a relative marker, and -ka- is some kind of tense marker.)

(13)

the red book	buku dzuku
the book is red	buku rakatsuku
the red books	mabuku matsuku
the books are red	mabuku akatsuku
the white book	buku ǰena
the book is white	buku rakačena
the white books	mabuku mačena
the books are white	mabuku akačena
an agreeable voice the voice is agreeable the voice that is agreeable	izwi rakatendeka

agreeable voices
the voices are agreeable } mazwi akatendeka
the voices that are agreeable

the broken bell
the bell is broken } ɗare rakaputsika
the bell that is broken

the broken bells
the bells **are** broken } matare akaputsika
the bells that are broken

the sharp hoe
the hoe is sharp } ɓadza rakapinza
the hoe that is sharp

the sharp hoes
the hoes are sharp } mapadza akapinza
the hoes that are sharp

In discussing this data, we will refer to adjectives such as -čena 'white' as "real adjectives," and those such as tendeka, 'agreeable,' as "predicate adjectives." (Although they may not function as predicate adjectives do in English, this term will be adopted from Heny, 1972.)

Our class 5 morphophonemic rule as formulated has no trouble explaining such forms. However, one might **also** wish to account for why some adjectives can never appear as real adjectives, but only as predicate adjectives. If one could argue that adjectives which occur only as predicate adjectives were underlyingly verbs, then this distribution could be accounted for, and the lack of stem-initial voicing of stops would be doubly explained: first, because the stem-initial stop of a predicate adjective is never in word-initial position, and second because the initial stop of a verb is never affected by our class 5 morphophonemic rule.

While there appears to be sufficient evidence to argue for the latter solution, this is not within the scope of our paper and we will not present the details of the argument. Were such an argument to be accepted, however, we would have more choice in formulating our rule. Instead of specifying [+5], we could choose to note [-V].

We **now** come to the question of implosives in our analysis. It appears that there are only two implosive consonants in Shona, ɗ and ɓ. There are thus no non-stop, no velar, and no voiceless implosives that appear in the phonetic representation of the language.

In classes 5 and 6, these voiced implosives alternate either with voiced implosives, e.g.,

(14)	class 5	class 6	
	ɓuri	maɓuri	'hole'
	ɗutu	maɗutu	'strong wind'

or with voiceless non-implosives.

(15)

class 5	class 6	
ɓipito	mapipito	'shoulder'
ɗura	matura	'granary'

In the cases where voiced implosives alternate with voiceless non-implosives we are faced with five basic alternatives to account for the alternation.

(16) a) [voiced implosive] > [voiceless implosive] > [voiceless non-implosive] in class 6

e.g., ɓ > ɸ > p

b) [voiced implosive] > [voiceless non-implosive] in class 6

e.g., ɓ > p

c) [voiceless implosive] > [voiced implosive] in class 5

e.g., ɸ > ɓ

d) [voiceless non-implosive] > [voiced non-implosive] > [voiced implosive] in class 5

e.g., p > b > ɓ

e) [voiceless non-implosive] > [voiced implosive] in class 5

e.g., p > ɓ

Alternatives a and b are ruled out becuase, as we have argued before, to devoice in class 6 rather than voice in class 5 puts an added weight on the grammar of explaining why all class 5 stop-initial words begin with a voiced segment and under just what circumstances the stops devoice, everything being much more predictable under our analysis of voicing in class 5. Alternatives a and c, whereby we come about the alternation via a step including a voiceless implosive are a priori infeasible and ruled out because there are no such phonetic entities as voiceless implosives, and we prefer not to have any level in a derivation containing an unpronounceable item. Since we can avoid the postulation of so abstract a phoneme as a voiceless implosive, we have chosen to do so. This is not, however, to say that we are totally against the use of abstract entities

in a solution.

Of the two remaining alternatives under our analysis,

d) p > b > ɓ
e) p > ɓ

it appears that d would be the most attractive, in that the voicing and implosivizing are broken up into two clear steps rather than voiceless non-implosives becoming voiced implosives in one fell swoop. However, if one were to simply state that all instances of non-implosive /b/ and /d/ become implosive, i.e. something like

(17) $\begin{bmatrix} C \\ +stop \\ -impl \\ -vel \\ +vc \end{bmatrix} > [+impl]$

as a general clean-up rule for,

(12) $\begin{bmatrix} C \\ +stop \\ -vc \\ +5 \end{bmatrix} > [+vc]/\#\_\_\_\_\_$

we appear to run into trouble with borrowed words where /b/ and /d/ are never implosivized.[7]

(18)

class 5	class 6	
biza	mabiza	'horse'
damu	madamu	'dam'
dada	madada	'duck'

It thus appears that the proposed implosivizing rule does not apply to the entire lexicon of the language, but rather only to native Shona vocabulary. We are not splitting the vocabulary in half just at whim, however. The native Shona lexicon also differs from the borrowed vocabulary in that borrowed words never undergo our proposed morphophonemic voicing rule, e.g.,

(19)

class 5	class 6	
ketane	maketane	'chain'
peni	mapeni	'penny'
purisa	mapurisa	'police'

Thus it appears that not only our morphophonemic voicing rule, but also the implosivizing rule, must be "marked" as applying only to native Shona vocabulary. In other words, we are saying that at least some rules must be sensitive to a $[\pm$ native$]$ feature. Our rules now look like the following.

(3) $\begin{bmatrix} C \\ +stop \\ -vc \\ +5 \\ +native \end{bmatrix} > [+vc]/\#\_\_\_\_\_$

(20) $$\begin{bmatrix} C \\ +\text{stop} \\ +\text{vc} \\ -\text{velar} \\ +\text{native} \\ -\text{delayed release} \end{bmatrix} > [+\text{impl}]$$

(Note that none of the other analyses of voicing in class 5 make any claims about the state of implosives whatsoever.)

We also might, although it is not integral to do so, posit an underlying implosive /ɗ/ for native Shona, even though it is abstract in that it never shows up on the surface. An advantage and reason for doing this is to give more regularity to the underlying phonemic system in respect to the set of implosives. That is to say, under such an analysis all native Shona voiced stops would be underlying implosives rather than just the non-velar ones. As such, the final version of our implosivizing rule would be,

(21) $$\begin{bmatrix} C \\ +\text{stop} \\ -\text{vc} \\ -\text{impl} \\ +\text{native} \end{bmatrix} > [+\text{impl}]$$

with a general rule that de-implosivizes all velar implosives.

(22) $$\begin{bmatrix} C \\ +\text{velar} \\ +\text{impl} \end{bmatrix} > [-\text{impl}]$$

Thus far, we have given the impression that our native Shona morphophonemic voicing rule as it stands is exceptionless. Unfortunately, this is not the case, because we have forms like the following,

(23)

class 5	class 6	
tsime	matsime	'well'
pfumo	mapfumo	'spear'

where we would expect voicing in class 5 but do not get it.

In the cases of the affricates of class 5 we seem to be encroaching upon another morphophonemic process of sorts whereby forms like the following are acceptable.

(24)

class 5	
dzime	'big well'
bvumo	'big spear'

A voicing of the stops is permissible to give the noun the added dimension of ugliness, bigness, clumsiness, or some pejorative quality of this nature, although the extent of this process is not yet clear to us. And

this happens not only in classes 5 and 6, but also, for example, in class 21.

(25)

class 21	
komana	'child'
zigomana	'big child'

If these initial segments voiced in class 5, this would lead to ambiguity; that is, dzime could mean either 'well,' or 'big well.' It appears that this process, whatever the full ramifications of it are, is not reversible as such in that the devoicing of a regularly voiced stop does not give us a less ugly, large, or clumsy connotation, and in fact gives us ungrammatical forms.

(26)

bveni / mapfeni
but, *pfeni

bvupa / mapfupa
but, *pfupa

An amazing fact about all of this is that only pf and ts are ever exceptions to our morphophonemic rule of voicing in class 5. For this reason, these and only these phonemes may be overtly affected by the pejorativization rule. That is to say, /p/, /t/, /k/, /č/ are never overtly affected by this pejorativization rule, as they are always voiced in class 5 by the morphophonemic rule.

One analysis that could bring some sort of order to this chaos involving the class 5 and 6 affricates ts and pf is to claim that those root-initial affricates that show the voicing alternation and are normally voiced in class 5 are also inherently big, ugly, clumsy, or whatever; those that are inherently voiceless are not and only acquire voicing when a special quality of largeness or the like is attributed to them. Under such an analysis, it is proposed that our morphophonemic voicing rule never comes into play with ts and pf.

Such facts seem to be borne out to some extent, because we have

(27)

class 5	class 6	
bvembe	mapfembe	'watermelon'
bvupa	mapfupa	'bone'
bveni	mapfeni	'baboon'
dzimbo	matsimbo	'suffering'
dziwa	matsiwa	'mucus'

where the affricate voices, as opposed to class 5 pfumo 'spear', tsime 'footprint', pfuro 'pond', where the affricate does not voice. That is to say, to call a baboon ugly is to carry coals to Newcastle and the pf of pfeni thus is naturally voiced, whereas semantic forms like 'pond' and 'footprint' are more innocuous and only acquire such pejorative meanings specially.

In claims whereby the affricates ts and pf are never affected by our rule of morphophonemic voicing, one must block it from applying in some fashion. The simplest manner would seem to be as follows.

(28) /#_____V

In such a claim, one might be implying that ts and pf consist of two different phonemes each, whereas /ǰ/ and /č/ are single phonemes and undergo the rule as if they were just one phoneme. However, to make such a claim that ts and pf never undergo our morphophonemic voicing rule, along with the claims, implications, and necessary finaglings needed to show that there is a semantic division in the lexicon corresponding to the division between those ts and pf that voice in class 5 and those that do not, is unsupportable for several reasons, among which are the following.

(29) a) There is no independent evidence to support the claim that ts, dz, pf, and bv are of differend phoneme status from č and ǰ.

b) There is no natural phonological reason why the conditioning environment for our voicing rule should be /#____V instead of merely /#_____.

c) The semantic differentiation between "normally" voiced affricate-initial words which alternate with a voiceless affricate is not seen as cold hard fact but only as a tendency of sorts on the basis of the limited forms available.

As such, we are unfortunately left to mark those affricates ts and pf that, in a handful of forms, do not voice root-initially in class 5, as exceptions to our morphophonemic rule and seek perhaps an historical explanation for the peculiar phenomena surrounding these two phonetic sequences.[8]

Our discussion, and the formulation of our rule, have led the reader to believe that the voicing of root-initial stops in class 5 occurs only when those stops are also word-initial. There appear to be two exceptions to this, however, and examples of each follow.

(30)	class 5	class 6	
	ɓaŋga	mapaŋga	'knife'
	zibaŋga	mazipaŋga	'huge knife'
	ɗaŋga	mataŋga	'corral'
	ziɗaŋga	mazitaŋga	'huge corral'
	bveni	mapfeni	'baboon'
	zibveni	mazipfeni	'huge baboon'
	bvupa	mapfupa	'bone'
	zibvupa	mazipfupa	'huge bone'

ganda ibvawi	makanda mapfawi	'the skin is tender'
ǰaya iɓenzi	maǰaya mapenzi	'the young man is crazy'

There are several possible explanations for both of these apparent counter-examples to our claim, and we shall examine each alternative in turn, eventually choosing one as being more probable than the others.

The first and most obvious way of accounting for the above forms is to say that the affixes #i- and -zi- are exceptions to our proposed rule in that they allow voicing of a root-initial stop that is not also word-initial. According to this solution, then, these prefixes would have to be marked as being able to optionally intervene, i.e.,

(31) $\#\begin{pmatrix} \text{i-} \\ \text{zi-} \end{pmatrix}$ _____

While this analysis works--there are, after all, no other affixes which allow root-initial stop voicing in class 5--we will eventually discard marking them as exceptions in favor of another solution.

In examining the above forms which contain #i-, we see two possible explanations for the data other than marking #i- as an exception. The first possibility is to claim that our rule may be altered so that root-initial voicing could take place even if that root-initial stop were not word-initial, but were preceded by an optional vowel. Under such a solution, our rule would look like the following.

(32) $\begin{bmatrix} \text{C} \\ \text{+stop} \\ \text{-vc} \\ \text{+5} \\ \text{+native} \end{bmatrix} > [\text{+vc}] / \#(\text{V})$ _____

We do not accept this reformulation, however, for it appears to be an ad hoc way of accounting for the data and has virtually no natural phonological explanation. Since there seem to be no other prefixes which consist of a single vowel, this change in our proposed rule cannot be supported by any other forms and is merely a way of "doctoring it up."

The third possible explanation for why root-initial stops voice even when they are preceded by the prefix #i- is the analysis we will accept as being preferable to the others. Notice that when #i- is prefixed onto a noun, the result is always a sentential element. We will argue, therefore, that the #i- which is prefixed onto these nouns is <u>not</u> to be considered to be a prefix in the generally-accepted sense of that word, but that the element should be analyzed as a word. In this manner, the underlying form for <u>iɓadza</u> would appear as <u>#i#padza</u> and not as *<u>#i+padza</u>. Our rule of voicing stops in word-initial position in class 5 would then not be blocked from applying, and will produce the correct surface form, <u>iɓadza</u>.

Earlier in the paper proof was offered of a rule that voices initial stops to change the meaning of a word (phonemic as opposed to our non-phonemic **voi**cing). In such cases, the voiced alternant had the added meaning 'big, huge, ugly' or something similar. Now we see that there is an infix, -zi-, which also means 'big' when affixed to a noun, and appears to be an exception to our voicing rule, in that in class 5, the stem-initial stop voices even though it is not in word-initial position. Our explanation of this unexpected voicing of class 5 root-initial stops is the following: Since stem-initial voiced stops are obligatory when -zi- is added to forms that would otherwise be affected by our morpho-phonemic voicing rule, but voiceless stem-initial stops are obligatory when -zi- is added to forms not normally affected by our rule, the simplest method of accounting for the facts is to posit the insertion of -zi- 'big' at a synchronic level later than that of the noun class marker; otherwise, we would have no practical means of explaining why stops voice after #zi- (class 5) but not after #ma-zi- (class 6). Below is a sample derivation.

(33)

	class 5	class 6
u.f.	paŋga	paŋga
class marker	------	ma+paŋga
our rule	ɓaŋga	---------
-zi-	zi+ɓaŋga	ma+zi+paŋga
surface	ziɓaŋga	mazipaŋga

It has been shown that the two apparent counter-examples to our proposed rule for class 5,

(3) $$C \begin{bmatrix} +stop \\ -vc \\ +5 \\ +native \end{bmatrix} > [+vc]/\#\_\_\_\_\_$$

are really not counter-examples at all. The sentence-forming element #i- is probably a word in and of itself, and therefore the consonant immediately following it is word-initial and subject to our morpho-phonemic rule voicing initial stops in class 5. The infix -zi-, meaning 'big, huge', can co-occur with root-initial voiced stops because it is introduced **after** our rule of morphophonemic voicing in class 5 has applied.

We can offer yet another piece of evidence as to the validity of our claim that word-initial position is the only environment in which a stem-initial stop of a class 5 noun or adjective is subject to our voicing rule. There is a small class of nouns (peŋgo 'insane', tendi 'believing', dzidzi 'student', fumi 'rich', rombo 'poor', fara 'sickly') that, when used to modify a class 5 noun, show a #mu- prefix and the stem-initial stop is not affected by our class 5 voicing rule.

(34)

class 5	class 6	
ǰaya mupeŋgo	maǰaya mapeŋgo	'insane young man'
ǰaya mutendi	maǰaya matendi	'believing young man'
ǰaya mudzidzi	maǰaya madzidzi	'student young man'

This prefix #mu- appears to be a marker of personification, even though non-human nouns can also appear in these constructions:

(35) class 5: ruwa mufara
6: maruwa mafara 'sickly flower'

class 9: nika mufara
10: nika mifara 'sickly country'

Note, however, that if we speak of a rich country,

(36) class 9: nika fumi (*nika mufumi)
10: kanika kafumi

we are not personifying, and therefore we do not use #mu-. This is evidence that the #mu- of class 5 is not a "fixed" prefix for any certain adjectives.

Until now, all nouns in class 5 have been affected by our morphophonemic rule of voicing, because their stem-initial stops also happened to be word-initial. Now we have evidence that nouns in class 5 whose stem-initial stop is not word-initial are not affected by our voicing rule for class 5 nouns and adjectives. Therefore, it is imperative that we include word-initial position in the formulation of the rule.

To summarize, we have seen that our analysis of the voicing phenomenon in class 5 is by far preferable to the other analyses proposed to date, because we are able to account for the peculiarities surrounding the voicing in class 5 in a natural phonological way, something that has been impossible under the proposals of Fivaz, Heny, etc. That is to say, we have shown that the voicing phenomenon is due to a morphophonemic process which is limited to word-initial stops in class 5, rather than to a purely morphological "prefixing" process, as the other analyses would have it. Within such a standpoint, a prefixing process, whatever the exact proposed ramifications are, fails to explain why this "prefixing" only affects some root-initial consonants some of the time and vowels not at all, whereas all other noun class prefixes of both series affect, or are present with, all root-initial segments all of the time.

We have also delved to some extent into the status of implosives in the language, the existence of a "pejorativizing" root-initial segment voicing rule, the necessity of allowing phonological rules to be sensitive to the [± native] feature in Shona, and the like. Above all, we have seen that there is no prefix in the class 5 noun/adjective series, and that the voicing alternations that occur in classes 5 and 6 are due to the following morphophonemic rule.

(3) $$\begin{bmatrix} C \\ +stop \\ -vc \\ +5 \\ +native \end{bmatrix} \rightarrow [+vc] / \#\_\_\_\_\_$$

NOTES

1
This paper was presented at the Ninth Annual African Linguistics Conference held in East Lansing, Michigan, in 1978. We would like to thank our language consultant, Kokerai Rugara, for his patience and help, and Dr. Charles Kisseberth for his comments on an earlier version of this paper. Apecial thanks go to Lee A. Becker, for the time he gave us in revising our earlier paper.

2
6, ɗ represent implosives, b, d, etc., represent non-implosives.

3
Since tones in Shona have no bearing on our analysis, we will not mark our examples for tone.

4
Note that although there is a tendency for borrowed words with root-initial voiced stops to enter into classes 5 and 6, but those with voiceless stops to fall into class 9,

class 5/6	
damu/ madamu	'dam'
daramaǰi/ madaramaǰi	'dynamite'
buku/ mabuku	'book'
biza/ mabiza	'horse'
bara/ mabara	'wheel barrow'
ǰuzi/ maǰuzi	'jersey'

class 9	
tafura	'table'
tii	'tea'
kati	'coffee'
pen	'pen'
kapu	'cup'

this is not always the case, e.g.,

class 5/6	
purisa/ mapurisa	'police'
peni/ mapeni	'penny'

Since this tendency could be looked upon as being either the result of only voiced stop-initial words going into class 5, OR as the result of analogy to the end product of our rule of morphophonemic voicing, this is a moot point.

5
The class 9 "prefix" comes from /ni/ historically.

6
This i- is prefixed onto mono-syllabic words.

7
All cases of /b/ and /d/ in native Shona vocabulary are phonetically realized as implosives ($\ɓ,ɗ$).

8
It would be interesting, of course, if the historical explanation has the pejorative connotation rule come into play.

REFERENCES

Biehler, E. 1950. A Shona dictionary. Cape Town: Longmans, Green and Co., Ltd.
Chomsky, Noam, and Halle, Morris. 1968. The sound pattern of English. New York: Harper and Row, Publishers.
Doke, Clement M. 1931. A comparative study in Shona phonetics. Johannesburg: University of the Wirwarsrand.
Fivaz, Derek . 1970. Shona morphophonemics and morphosyntax. Johannesburg: University of the Wirwarsrand.
Fivaz, Derek, and Ratzlaff, Jeanette. 1969. Shona language lessons. Salisbury: Word of Life Publications.
Fortune, G. 1955. An analytical grammar of Shona. London: Longmans, Green and Co.
Gregersen, Edgar A. 1972. Prefix and pronoun in Bantu. International Journal of American Linguistics 33:3 Part 2 (supplement), 1-69.
Hannan, M. (compiler). 1974. Standard Shona dictionary, second edition. Salisbury: Rhodesia Literature Bureau.
Heny, Frank. 1972. Bantu lexical classes and semantic universals (with some remarks on how not to write phonological rules). Studies in African Linguistics 3.207-258.
Welmers, William E. 1973. African language structures. Berkeley: University of California Press.

Studies in the Linguistic Sciences
Volume 8, Number 1, Spring 1978

LINGUISTIC CONVERGENCE: INDO-ARYANIZATION OF DRAVIDIAN LANGUAGES

S. N. Sridhar

It is well-known that over three millennia of language contact on the Indian subcontinent has resulted in the convergence of linguistic features of the four language families of the area. A major part of this ongoing process is the adaptation of Indo-Aryan characteristics by the twenty-three or so Dravidian languages. The present paper is an attempt to study the various aspects of this phenomenon in broad outline. Following the introduction, the language contact situation between Sanskrit, Prakrit, and the modern Indo-Aryan languages on the one hand, and the literary and non-literary Dravidian languages on the other, is described. Next, variation in the nature and extent of Indo-Aryanization both within and across the Dravidian languages is outlined. Section four deals with the impact of Indo-Aryan on Dravidian phonology, morphology, lexicon and syntax; this is followed by a discussion of the processes of nativization of borrowed elements. Section six describes the sociolinguistic implications of Indo-Aryanization, especially the emergence of the manipravala style, the reinforcement of caste dialects and diglossia, and language attitudes. The conclusion raises some issues concerning the role of Indo-Aryanization in the evolution of Dravidian languages.

1. Introduction

It is often remarked that the Indian subcontinent has a genius for assimilating foreign influences without losing its essential character.* Nowhere is this resilience demonstrated more clearly than in the way the Dravidian languages have absorbed what, by any count, must be regarded as massive Indo-Aryan influence and yet retained their essential Dravidian character.[1]

In discussing the Indo-Aryanization of Dravidian languages, we are talking about an ongoing phenomenon that must have started about 3,500 years ago. The presence of linguistic features of possible Dravidian origin in the Rig Veda[2] (e.g., retroflex consonants) suggests Aryan contact with the Dravidian-speaking people as early as about 1500 B.C. This possibility is confirmed by the geographical distribution of the Dravidian languages--e.g., Brahui in Balochistan, Gondi in Madhya Pradesh, Malto in Bihar-West Bengal, among others. Ignoring the possibility of reverse-migration, this scatter suggests the presence of Dravidian-speaking people over most of India at one time, and their gradual recession before the advancing Aryans. However, our recorded history of Dravidian contact with Indo-Aryan goes back only to the early centuries B.C.

The present topic may be--and has been--studied from many points of view.[3] In the past, the main concentration has been the phonetic and

semantic modifications undergone by Indo-Aryan words in Dravidian languages. However, there is much more to the topic than that. In the present paper, my aim is to present an overall picture of the many facets of Indo-Aryanization of Dravidian. Following the introduction, section 2 discusses the reasons, sources and mechanisms of Indo-Aryan borrowings; section 3 outlines the variation in the extent of Indo-Aryanization; section 4 is a survey of the Indo-Aryan influence on Dravidian on various levels; section 5 is devoted to nativization; section 6 discusses the sociolinguistic implication of Indo-Aryanization, and section 7 comprises the conclusion.

Before proceeding further, I would like to point out that what follows is more in the nature of a bird's eye view than a microscopic examination of minutiae. Due to the paucity of materials and my own inexperience, I have not been able to do equal justice to all the languages involved. Yet, I hope the following discussion may serve at least as a programmatic sketch for a much needed full-length treatment of this important topic.

2. THE REASONS AND SOURCES OF BORROWING

The reasons for the prodigious influx of Indo-Aryan features into Dravidian are fairly well-known. First, Sanskrit has traditionally occupied an exalted position as the language of the elite in India, a position analogous to that of Latin in Europe until a few centuries ago. Sanskrit was not only the language of the scriptures, the epics and classical literature, it was also the repository of a vast body of philosophical, scientific and technical literature in every field. It enjoyed royal patronage, and the prevalence of Sanskrit-based education largely under the control of the Brahmans reinforced its prestige. All this led to the creation of a vast body of literature in the Dravidian languages, composed in conscious imitation of the Sanskrit models, bringing in its wake a veritable flood of Sanskrit words. While many of these words were introduced to meet the technical needs of specialized language-types (or "registers", for example, in grammar) and often to suit the demands of intricate versification (for example, ādiprāsa and antyaprāsa), a large number must have been introduced simply because it was prestigious to do so.

Secondly, Prakrit (and to a lesser extent, Pāli) played a crucial role in the Indo-Aryanization process. In some languages (e.g., Tamil), the words that came in via Prakrit far outnumber those borrowed directly from Sanskrit. There are two major reasons for this Prakrit influence. It was the court language during some periods in the Dravidian South, for example, during the Sātavāhana and the early Pallava periods. Also, the South, especially Karanataka, was the haven of refuge for the Jainas and a large number of very prominent authors in the Dravidian languages were Jainas who often composed in Prakrit as well (e.g., Nēmichandra). Many prominent Prakrit authors also wrote in Dravidian languages, for example, Pushpadanta, Trivikrama, and Sakāṭāyana, who wrote in Kannada. Many of Indo-Aryan loan-words in the non-literary languages are also traceable to Prakrit forms.

The next major category of Indo-Aryan loanwords in Dravidian comes

from modern Indo-Aryan, especially Hindi-Urdu and Marathi in the literary languages of the South, Oriya in Pengo and Kui, Sindhi in Brahui, Marathi in Kolami, Hindi in Kurukh, etc.

Parts of South India came under Muslim rule in various periods, especially during the reign of the Bahamani kings in Andhra and North Karnataka, and that of Hyder Ali and Tippu Sultan in South Karnataka. During the latter period, Persian was the official language of the princely state of Mysore, with official records being kept in that language. The prolonged Muslim rule has contributed a large number of words in the fieldsof law, administration,and land-management, as well as those relating to Muslim culture (see Sridhar, 1975). Finally, the increased commerce between the North and South since the beginning of the Independence movement has resulted in the introduction of more Hindi words (e.g., kālāpānī, gherāo, bandh, čarkhā, etc.).

As for the influence of Marathi, both geography and politics have been contributing factors. Maharashtra shares borders with both Karnataka and Andhra Pradesh, and there is intensive code-mixing in these border areas (see, Gumperz and Wilson, 1971; Upadhyaya, 1971). Politically, much of North Karnataka came under the Bombay Presidency until the reorganization of the states in 1956, and the administrative language as well as the language of education in this region was Marathi. Further South, the Tanjavur area in Tamil Nadu was under Maratha rule for nearly two centuries until 1800 A.D.(Subramonian and Ganesha Sundaram, 1954). The Deśastha Mādhwa brahmins in the South have traditionally been bilingual in Marathi and Kannada, Telugu or Tamil.

The non-literary languages of the South (Toḍa, Koṭa, Kodagu, Baḍaga, etc.) have also numerous Indo-Aryan borrowings in them, but these seem to have resulted from diffusion from their literary neighbors (Emeneau and Burrow, 1962). As for the non-literary languages outside the major Dravidian belt, their language contact with Indo-Aryan needs no explanation. Speakers of these languages often indulge in such occupations as wood-gathering, basket-weaving, and perform menial jobs, and in general are crucially dependent on their Indo-Aryan neighbors for economic survival. Bilingualism in the dominant language of the region is for them a cardinal necessity. Under these circumstances, it comes as no surprise that some of these languages show much more radical changes in response to Indo-Aryan influence than their literary sisters.

The preceding discussion assumes, following the long recognized theory, that the main agency responsible for the transfer of traits in a language-contact situation is the bilingual individual. Several recentsociolinguistic studies of language use in bilingual communities (e.g., Gumperz and Wilson, 1971, and Upadhyaya, 1971, in the Maharashtra-Karnataka border) give us an insight into the processes by which the present day convergence of features in Indian language must have come about. In particular, Kachru (1978) argues that code-mixing may be a crucial factor in the kind of language change we have been discussing.

3. VARIATION IN THE EXTENT OF INDO-ARYANIZATION

The Dravidian languages show a great deal of inter-language and intra-language variation in the extent of Indo-Aryan influence. The non-literary languages spoken in the Indo-Aryan region seem to have undergone a greater amount of structural (morphological, syntactic) influence (see section 4.3), whereas the literary languages of the South have been influenced predominently on the lexical level. Second, the non-literary languages contain few unassimilated Sanskrit lexical items while this category plays an important role in the lexicons of the literary languages. Third, it is suggested (Andronov, 1964) that the dialectal dissolution of Dravidian was already well under way before Indo-Aryan began to exert its influence. If this is correct, [4] then that fact could account for some of the variation. Finally, the extent to which a given language has been influenced by Indo-Aryan also depends on the language attitudes of its speakers. For example, Tamil has traditionally been more conservative than its neighbors in this regard--a fact commonly attributed to the fierce loyalty of the Tamils to their language and their desire to preserve its "purity" (see section 6.3 for a detailed discussion). In contrast with this, consider Bhattacharya's observation that speakers of Ollari (a non-literary Dravidian language spoken in the Koraput district of Orissa) disclaim proficiency in their mother tongue (Bhattacharya, 1957).

There is also a great deal of intra-language variation with regard to the extent of Indo-Aryan influence. In this respect, the chief variables involved are (i) style, (ii) caste, and (iii) topic of discourse.

Dravidian languages show more Indo-Aryan features in written style than in speech, and within the written style, more in poetry than in prose. It is well known that during the various periods in the history of Dravidian literature, a style of writing characterized by heavy Sanskrit borrowings, often together with the original Sanskrit declensions and long compounds (e.g., Te. dharmanivahanōdyōgamummanaku) became fashionable (see Kunjunni Raja, 1972). This style, commonly referred to as Maṇipravāḷa (combination of gems and corals) reached its peak in Malayalam (e.g., Chandrōtsava), but it is met with in the other three literary languages as well (for example, the inscriptions of Rāja Rāja Chōla and in the Jaina and Vaiṣṇava philosophical literature in Tamil, in the work of Nannaya and Potana in Telugu, and to give a modern example, Kuvempu's Rāmāyaṇadarśana in Kannada). Like the epic sytle of Milton, Maṇipravāḷa never found universal favor and is largely a literary and linguistic curiosity now, being used marginally in orthodox Brahman doctrinaire tracts, if at all.

The extent of Indo-Aryan features also varies according to the caste of the speaker (writer). Bright (1966), Bright and Ramanujan (1964), and Ramanujan (1968), among others, have observed that one of the variables distinguishing the Brahman dialects[5] in Tamil and Kannada from the non-Brahman is the presence of Sanskrit phonemes and clusters and vocabulary. Here we see what the anthropologist Srinivas calls 'Sanskritization' (Srinivas, 1966) operating in a literal sense, the Sanskrit features serving as a symbol of caste identity and distinctness.

The emergence of hypercorrect forms with unnecessary aspiration may also be regarded as a manifestation of this phenomenon.

The third parameter of variation is the topic of discourse: Indo-Aryan features are more likely to be present in the discussion of some topics than in others. For instance, religious discourses, in particular, explication of scriptures, known as upayāsas in the South, contain heavy Sanskrit borrowings, as do texts and oral discourses on poetics, aesthetics and philosophy, as well as those on the social sciences in general. By the same token, one finds extensive use of Hindi-Urdu words (ultimately of Perso-Arabic origin) in discussions of law, land-revenue, Hindustani (or North Indian) music, wrestling, etc. (for a detailed discussion of this register-oriented code-mixing, see Sridhar, 1975 and 1978). In modern times, Sanskrit has been serving as the principal source of new coinings in all the scientific and technical fields. (Tamil is, once again, an exception to this statement, the emphasis there being on the exploitation of native resources.)

The foregoing discussion of variation in the extent of Indo-Aryanization should suffice to make clear that any facile generalization about Indo-Aryanization of Dravidian languages, and especially the non-literary ones, is likely to prove wrong. Here, one can only note some general tendencies, always keeping in mind the complex inter- and intra-language variation. With this caveat, let us now turn to an examination of the influence of Indo-Aryan on various levels.

4. INDO-ARYANIZATION ON VARIOUS LEVELS

4.1 Vocabulary

Anyone who looks at the dictionaries of Dravidian languages cannot but be struck by the extravagant proportion of Indo-Aryan words in general, and Sanskrit words in particular, contained in them. Emeneau and Burrow (1962) rightly attribute this to "the tendency for all four of the Dravidian literary languages in the South to make literary use of the total Sanskrit lexicon indiscriminately (p.1)." Even in the colloquial language, so many words of Indo-Aryan origin have found their way and been assimilated that it is hard to speak even a few sentences without using some of these words. It was no doubt partly because of this high proportion of Sanskritic elements (and the similarity in the sound values of the graphemes in the scripts [counting the grantha script in Tamil]) that, until just over a century ago, it was generally believed that the Dravidian languages were descended from Sanskrit. Even Western scholars like Pope, Colebrooke, Carey, and Wilkins subscribed to this view. Caldwell (1903) spends a considerable part of his classic work dispelling this misconception.

It is impossible to give a list of borrowings in Dravidian (for a partial list,[5] see Emeneau and Burrow, 1962) or even to specify the conceptual fields they denote, for there is no field of activity which has not been influenced by them. In many cases, the native terms have been replaced by the borrowed items, even in the areas generally regarded as part of the basic vocabulary. A case in point is the set of direction terms in Kannaḍa, where the native words mūḍal, paḍuvaṇ, baḍagal, and tenkal have been replaced by the Sanskrit

pūrva, paścima, uttara, dakṣiṇa, even in everyday colloquial use. Sjoberg and Sjoberg (1956) estimate that about twenty percent of the "non-cultural" part of the basic vocabulary in literary Dravidian now consists of loan-words from Indo-Aryan, creating problems for glottochronology (for a detailed discussion, see Sjoberg and Sjoberg, 1956, and Sreekantaiya, 1956). The proportion of borrowings in those areas that are more prone to cultural influence is, obviously, much higher, probably as much as fifty to sixty percent. Among the non-literary languages, Emeneau (1962a:20) estimates that the vocabulary of Kolami (Wardha dialect) contains approximately thirty-five percent of words with Marathi etymologies. The case of other non-literary languages is perhaps not much different.

Certain salient features of the Indo-Aryan loan-words in Dravidian may be noted. First, the loans are not restricted to any one register or set of registers, they are everywhere--in basic as well as technical vocabulary, in the written and the spoken language, in the language of intimate conversations and personal letters as well as that of newspapers and public speeches. Second, the loans are not restricted to content words alone, but include a large number of function words such as quantifiers (kēvala, bahaḷa, alpa), intensifiers (atyanta, mātra). adverbial conjunctions (akasmāt, bahuśah), etc., as well as bound morphemes of various kinds (see next section). Third, a large number of Sanskrit words have been borrowed both in their root form (e.g., jiva) and in their derived forms (e.g., jīvanta, śaktiyuta, aprakaṭita, tāratamya, etc.). Fourth, the Dravidian languages, especially the South Dravidian languages, have a substantial number of Indo-Aryan loan words in common (e.g., nadi 'river', candra 'moon', rātri 'night', bhūmi 'earth', etc.) and these words often undergo similar phonetic and semantic changes in these languages (e.g., anumāna(m)(u) 'suspicion' [<Skt. id. 'inference, guess'], see section 4.5 below). Fifth, all the literary languages have sets of two forms, one taken from Sanskrit and employed with minimal modifi cation, and the other either borrowed from Prakrit or modified in the borrowing language, or both. Often these pairs are distributed in a diglossic fashion (see discussion in section 6.2). Finally, the majority of Indo-Aryan loans undergo various kinds of phonetic and semantic changes, some of which are common to all the Dravidian languages and some language-specific (see section 5 for more details).

4.2 Phonology

This large-scale infiltration of Indo-Aryan words has resulted in the introduction of several phonemes alien to the Dravidian phonological structure. Among these, the most important are the aspirated stops (kh, ch, ṭh, th, ph, and their voiced counterparts) as in, for example, Ka. khāra 'spicy hot', chatri 'umbrella', ṭhāṇe 'station', kathe 'story', phala 'fruit', and ghaṇṭe 'bell', 'hour' jhari 'centipede', mūdha 'fool', nidhāna 'slow', bhāra 'heavy'. Other new introductions are the distinction between s, ś, and ṣ as in saśeṣha 'yet to be finished', and the diphthong au as in gauḍa 'village chief'. The scripts of all the literary languages except Tamil now possess characters for the full complement of Sanskrit phonemes. Even in Tamil the grantha script accomodates these phonemes.

The Indo-Aryan loanwords have also resulted in the loss or violation of some positional constraints in Dravidian. Thus, the presence of voiced consonants in the word-initial position as in gamaka, initial consonant clusters and non-homorganic plosive clusters as in prārabdha, word-final obstruents, as in tākat (<H-U), among others, may be attributed to Indo-Aryan influence.

However, it is important to note that thesenon-native phonemes and consonant clusters are not found in all speakers and in all styles (see Sjoberg, 1962, and Fowler, 1954). As already noted, these features mark certain styles and their presence is determined by socio-linguistic variables (see section 6.1 for further discussion).

The non-literary languages have also adopted many phonological features from Indo-Aryan. For example, Brahui and Kurukh have developed nasal vowels of the Indo-Aryan type. Brahui has also lost the short e and o characteristic of Dravidian (Emeneau, 1962). Similarly, Pengo (another non-literary Dravidian language spoken in Orissa) has lost the distinction between j and z that it inherited from Proto-Central Dravidian, presumably under the influenee of Oriya, which lacks that distinction. Pengo has also developed the feature of aspiration, although it is confined mostly to loanwords from Oriya and used only sporadically even there (Burrow and Bhattacharya, 1970).

4.2 Morphology

On the morphological level, the most striking characteristic is the productive use of a large number of Indo-Aryan derivational affixes. Sanskrit suffixes such as -anta, -iṣṭha, -kāra, are freely used (e.g., Ka. hanavanta 'rich', Te. kōpiṣṭhi 'irascible', Ma. vasaḷkāram 'imprecation'), as are the Hindi-Urdu ones like -dār (Ka. pattēdāra 'detective') -khōr (te. lan̆cakōru 'corrupt man'), -vāla (Ka. mīsevāla 'the moustached man'), and -giri (Ka., Te. gumāstagiri 'clerk-ship').

The employment of prefixation as a productive work-forming device in Dravidian must be attributed to the Indo-Aryan impact and is seen in the number of forms involving, for example, Sanskrit apa-, ava-, ati-, su-, pari-, as well as the Hindi-Urdu prefixes, bē-,māji-,gair-, etc. Consider, for example, the following. Ka. apakīrti 'ignominy', Te. apa dūru 'insinuation', Ka., Te. aparūpa(m) 'rare', Ka. ati mātu 'garrulousness', Ka. sulōcana 'eyeglasses', Ka. pari vīkṣane 'survey', Ka. prati obba 'each one', as well as Ka., Te. bēniyat(tu) 'disloyalty', gairuhājari 'absence', māji adhyakṣa 'ex-president', etc. Most of these affixes freely collocate with native roots as well, although it is more common to attach them to other borrowed words and use them in a sense not attested in the source language (as in the case os the Ka. pattēdāra.)

In a different vein, Andronov (1964) has observed that the gradual loss of many negative-incorporated forms in the major Dravidian languages and the use of analytical modes of expressing negation by means of a positive verb followed by a negative word may be due to the influence of Indo-Aryan. Thus, in Kannada, for example, forms such as hōga, māḍa have gone out of use, having been replaced by hōgōd-illa and māḍōd-illa.

Turning to the non-literary languages, Bloch (1954) has pointed out that the infinitive suffix -na in Gondi is indubitably borrowed from Hindi-Urdu. Another infinitive suffix -le in the same language is attributed to the influence of Eastern Marathi.

Pengo occasionally employs the Oriya genitive morpheme in -r as in rājar kuvar 'king's courtyard', and kumar ar 'potter's wife'. In Pengo, one also finds a marginal occurrence of the Oriya locative suffix -ne as in barhane 'in the rain'. More regular and widespread is the use of a number of Oriya post-positions, such as the instrumental hudun, san, 'with', the dative kajin 'for the sake of', and the locative bitre 'in', tarento 'beneath' (Burrow and Bhattacharya, 1970,41 ff.).

Emeneau (1962b) has brought to light an interesting case of morphological borrowing in Brahui, namely the allomorphs kar and kam ('to do, make') from Sindhi and Balochi, respectively, on the lines of the established Brahui irregular structure with -nm and -r. Brahui has undergone several other significant morphological changes not attested elsewhere in Dravidian, but these changes, as Emeneau rightly suggests, are probably directly traceable to the influence of Balochi rather than of the Indo-Aryan Sindhi.

4.3 Syntax

The influence of Indo-Aryan is least pronounced on the level of syntax. Although a number of syntactic patterns modelled on those of Sanskrit were introduced by various authors from time to time,[7] by and large, few syntactic traits in contemporary Dravidian can be traced to Indo-Aryan influence. There are a few notable exceptions, however. Modes of compounding such as the avyayībhāva (where the first component carries the major semantic burden , e.g., yathā śakti 'according to one's ability') seem to go against the general Dravidian tendency to place the determiner before the determined and are probably due to the influence of Sanskrit. Similarly, the passive construction, (for example, Ka. huḍugiyinda bāgilu tereyalpaṭṭitu 'girl- instr.-door-to open (inf.)-experienced' "the door was opened by the girl"), seldom used as it is, is also likely to have developed under the influence of Sanskrit, a tendency further reinforced by the influence of English. The third major syntactic trait in Dravidian traceable to the impact of Indo-Aryan is the clausal mode of relative-clause formation, (described below), which is widely attested in all the literary Dravidian languages.

As is well-known, the preferred mode of relative clause formation in Dravidian is by turning the main verb into a participle and deleting the coreferential noun, as in Ka. hasiru sīre uṭṭiruva hengasu nanna heṇḍati 'green-saree-wearing-woman (is) my-wife'. In addition to this construction, we now find another construction, closely modelled on the Indo-Aryan, e.g., Ka., yāva hengasu hasiru sīre uṭṭiddāḷo avaḷu nanna heṇḍati, 'which-woman-green-saree-is-wearing+ indefinitizer-she (is) my-wife'. "The woman who is wearing the green saree is my wife." It may be noted that this construction presents one of the few cases in Dravidian where the main verb in a subordinate clause is not changed into a non-finite form, and a coreferential subject is not deleted. Nadkarni (1970) convincingly argues for the Indo-

Aryan origin of this construction.

The non-literary languages in day-to-day contact with Indo-Aryan show greater receptivity to syntactic influence. For example, Brahui has developed enclitic pronouns which occur suffixed to verbs and nouns (Emeneau, 1962a:64). Brahui has also lost the inherited Dravidian system of natural gender in the demonstrative pronouns. Both of these traits are, once again, probably due to Balochi influence. However, a third trait, the use of ki as a complementizer may be attributed as much to Sindhi as Balochi influence. Pengo has borrowed the morpheme ki from Oriya, which is used as an interrogative particle appended to the end of the sentence, as in venvatay ki? 'did you not hear?' Ki is also used in the sense of 'or'. Pengo also employs the Oriya conjunctive particles ar, are ('and','again') (Burrow and Bhattacharya, 1970:61). Gondi has lost the relative participle typical of Dravidian and has adopted the Hindi relative pronoun (Caldwell, 1903:50).

In concluding this section of the influence of Indo-Aryan on different levels, it must once again be emphasized that the prece ding characterization is by no means intended to be exhaustive. It is merely illustrative of the kind of influence Indo-Aryan has had on Dravidian.

5.0 NATIVIZATION

There has been a long tradition, in the native grammatical literature, of distinguishing two broad classes of borrowings from Indo-Aryan, namely tatsama ('same as') or words which undergo little (final vowel or consonant) or no modification, and tadbhava ('derived from') or words which have been nativized or assimilated. For example, Kannada has both the unassimilated pakṣi (<Skt. pakṣin 'bird') and the assimilated hakki. In this context, "assimilation" refers only to phonological modification and does not include semantic changes that the Indo-Aryan loans undergo in Dravidian.

5.1 Phonetic Adaptations

Some sound changes in borrowed words are common to all four major languages. For example, the doubling of the final obstruent followed by the addition of the enunciative -u (e.g., sampattu 'wealth', hukummu 'order', tējassu 'halo'), the substitution of unaspirated stops for aspirated ones (e.g., parike 'broomstick'[<Skt. parikhā], the dropping of a final consonant (e.g., rāja 'king'[<Skt. rājan]), the substitution of s for ś and ṣ (ka. hasu[<Skt. paśu 'cattle'],simplification of consonant clusters by anaptyxis (e.g., Ma. kāriya[<Skt. kārya 'work'], intervocalic voicing of voiceless plosives (Ta. vigaḍan[<Skt. vikaṭa), and so on. Many of the changes are, however, language specific. In Kannada, for example, the final -ā of loan words changes to -e (e.g., mātā>mate, ilākhā>ilākhe, etc.) while in Telugu, such a change occurs only marginally. Similarly, Malyaylam speakers substitute l and ḷ for the Sanskrit t(d) and ṭ(ḍ) in certain contexts, e.g., when followed by a plosive (tālparyam, taḷbhavam, vaśaḷkāram, khaḷgam, etc.), whereas Tamil substitutes an r (Aiyar, 1973). Similarly, the regular substitution in Tamil of voiceless consonants for the voiced ones in

Sanskrit (e.g., cinkāram < Skt. sringāra) occurs much less frequently in other languages.

By and large, most of the phonemic substitutions in Dravidian follow the regular phonological rules of the languages concerned. This adaptation, together with the fact that a majority of Indo-Aryan loans commonly used in Dravidian have come via Prakrit, makes some forms hardly recognizable as originating from Sanskrit. A case in point is Ka. habba 'festival' (Skt.parva) which comes from Pkt. pavva (or pabba?) and, with the mutation of v and b in Kannada, enters as pabba in old Kannada, and with the change of initial p to h which took place around the 10th century, becomes habba in Middle and Modern Kannada.

At this point, it is worth mentioning that this tendency to regularize loanwords to make them conform to the native phonological structure operates more strongly in Tamil than in the other Dravidian languages. As a result, Tamil has fewer unassimilated Sanskrit words than the others. At least two reasons may be advanced to account for this. One is the influential injunction in Tolkāppiam (Sutras 397,401-2) according to which only those Sanskrit words could be used which either had letters common to Tamil and Sanskrit or had been modified in such a way that they could readily be represented in the Tamil script. This rule was scrupulously adhered to for over 1600 years until the great influx of Sanskrit words (including tatsamas) during the Vijayanagar period (16th century). The second,related reason may have been spelling pronunciation, the Tamil script lacking characters for voiced consonants and apirates, and palatal and retroflex sibilants. Note that the other literary languages have been more lax in this regard and have moved in the opposite direction, extending their phonemic (and graphemic) inventory to accomodate the foreign sounds.

5.2 Semantic Changes

Many of the loan-words from Indo-Aryan undergo various kinds of meaning change when used in Dravidian. Some of these changes are common to South Dravidian as a whole, while others are innovations peculiar to one or two languages. The following table, based on Aiyar 1934-35, illustrates some common semantic developments.

Source Word	Meaning in Sanskrit	Meaning in Dravidian
anumānam	inference; guess	suspicion
antaḥstaḥ	in the middle state	story of a building; high rank.
asahya	unbearable	loathsome
avasthā	state	miserable condition
pāpam	sin, evil	sin,pity

Both Aiyar (op.cit) and Emeneau and Burrow (1962) contain many more examples of this phenomenon, traceable probably to diffusion within South Dravidian. Similar changes may be noticed in Hindi-Urdu borrowings as well, e.g., Ka. Te., khāyam (<id. 'stand erect') 'permanent', dubāra(i)(<id. 'twice') 'excess,wastage'.

On the other hand, cases in which the same Sanskrit word has developed different meanings in the Dravidian languages are more frequent. For example, ārambham in Telugu has the meanings, 'pride' and 'killing' (literary), as well as the original Sanskrit meaning of 'beginning'. The same word in Kannada means 'cultivation' as well as 'beginning'. Similarly, āmōdamu in Telugu means 'dalliance' as well as 'frangrance'; in Kannada, the word is used only in its Sanskrit meaning of 'enjoyment'; piramātam in Tamil means 'excessive; as well as 'excellence'; the derivative of the same source in Kannada, pramāda has the meaning of 'blunder'. The Malayalam word gōṣṭhi means 'prank,contortions, gesture', while in Kannada and Telugu, the same word has the original Sanskrit meaning of 'assembly'.

The semantic changes that take place in Indo-Aryan words when used in Dravidian comprise of virtually every known type of deviation discussed in the linguistic literature. There is, for example, restriction of meaning in S. Dr. udyōga 'employment; (<Skt. 'work'); extension in Ka. anubhāva 'mystical experience' (<Skt. ' '); specialization in Ka. jangama 'a Virasaiva saint' (<Skt. 'mobile'); amelioration in Te. ālaya 'temple' (<Skt. 'house'); deterioration in Ka. adhvāna, Te. adhvānamu, Ta. attuvānam (<Skt. adhvan 'path, way') 'destruction, disarray', and so on. Similarly, Hindi-Urdu and Marathi words also change their meaning in various ways, e.g., Ka. čādi (<Ma. čahad ' ') 'petty complaint, informing on someone', Te., Ka. ustād (<H-U ustādi 'teacher') 'wrestler, gymnastics teacher', Ta. attāvanai (Ma. aṭhāvana 'recollection') 'index, ledger', Te., Ka. saṟa (<H-U ṣara'condition') 'post script', etc.

6.0 SOCIOLINGUISTIC IMPLICATIONS OF INDO-ARYANIZATION

It was remarked earlier (section 3) that all the Dravidian languages are not uniformly Indo-Aryanized, nor are all varieties in a given language influenced to the same degree. This variation is sociolinguistically conditioned to a large extent, and, in turn, contributes to an already complex sociolinguistic situation in the Dravidian languages. In this section we shall briefly examine this aspect.

6.1 Caste Dialects

The existence of caste-related differences in Dravidian languages has long been recognized. In particular, sharp differences are noted between the Brahman and non-Brahman dialects on the levels of phonology, morphology, and lexicon. While many of these differences have nothing to do with borrowing (e.g., Kannada ide(B) versus aite(NB) for 'to be 3rd sing. neut.), quite a few of the variables involve elements introduced in the process of Indo-Aryanization, in particular, Sanskritization. The most important of

these variables is aspiration in loanwords from Sanskrit. Even in these secular days, it is not uncommon to find parents and teachers upbraiding an unaspirating urchin for "speaking like a śudra." Another important variable is the existence of separate sets of lexical items in the Brahman and non-Brahman dialects. As Aiyar (1932) has pointed out with reference to Tulu, words borrowed directly from Sanskrit are far more numerous in the Brahman dialects, while words borrowed or adapted from Middle Indo-Aryan (Jaina Prākrit or Pāli) are common to both Brahman and non-Brahman dialects (e.g., Tu. upavāsa (B) vs. nompu (NB) 'fast'; Ta. tīrthõ (B) vs. tanni (NB) 'water'). Even when the same Indo-Aryan word has been borrowed into both dialects, it undergoes different modifications: (Tu. sāvira (B) vs. sāra (NB) 'thousand'; Ka. bahaḷa (B) vs. bāḷa (NB) 'much, very').

The reasons for this conservatism[8] of Brahman dialects are not hard to find. The preservation of Sanskrit features in their speech gives Brahmans a social identity which their insignificant number denies them. As a byproduct of the elitist position of Sanskrit in India, this feature gives them prestige as well. Moreover, as Ramanujan (1968) correctly points out, part of the reason may also be the high premium placed in that community on the accurate rendering of the texts and their conservative outlook toward cultural innovations in general.

The symbolic value of Sanskritic features in the Dravidian sociolinguistic context has also led to the inevitable phenomenon of hypercorrection. Thus we find scores of native Dravidian words restructured to incorporate prestige features such as aspiration. Kannada offers many examples of this sort: bhatta 'paddy', chaḷi 'cold', jhari 'centepede', etc. (cf. also, Te. iddharu 'two persons'). These hypercorrections are manifestations, in a literal as well as in the sociological sense in which Srinivas (1966) employs the term, of Sanskritization.

6.2 Diglossia

Indo-Aryanization--in particular, Sanskritization--has also contributed to the age-old cleavage in Dravidian languages between the literary and the colloquial styles. In this context, the term'literary style' is used in its broad sense, to refer to all styles of institutionalized use of language including that of the public platform, and is not restricted to the written form. There have been a number of studies of diglossia in Dravidian (e.g., Caldwell, 1903:78ff; Shanmugam Pillai, 1960; Radhakrishna, 1971; Nayak, 1967), all of which point out that one of the distinguishing characteristics of the 'High' style is the preponderance of Sanskritic features. The 'High' style is characterized by the presence of non-Dravidian features such as aspiration, and the use of unassimilated borrowings, as well as structures such as the passive construction and the clausal mode of relative clause formation. This penchant for mixing non-native elements in the literary style gives it a rather artificial and stilted character, and when carried to an extreme, as in the Maṇipravāḷa style, often becomes unintelligible to the uninitiated layman. It is not surprising, therefore, that from time to time, there have been movements, in all

the literary languages, to bring the literary style closer to the language of the common man, and a recurrent theme in these programs has been a call for reducing the Sanskritic component in the literary style (see the next section).

The association of Sanskrit with the High style in Dravidian has resulted in an interesting phenomenon: the diglossic distribution of multiple derivatives on the borrowed words. Dravidian languages provide many instances of two forms of the same Indo-Aryan loanword being used in different senses, one, typically the unassimilated Sanskrit word, retaining its original meaning or undergoing amelioration, and the other, typically the form borrowed or adapted from Prakrit, undergoing deterioration. The follwoing examples from Kannada are relevant: samsthe 'organization'; sante 'weekly village market'; parisattu 'institution'; parse 'annual fair'; dhana 'wealth'; dana 'cattle'; ārya 'respected person, an Aryan'; ajja 'grandfather, an old man'; vandhye 'infertile woman'; banje 'id (pej)'.

In this connection, mention must also be made of the resort to Sanskrit terms for euphemisms in polite conversation and in writing. Thus, the preference for terms like jaghana 'thighs', kuca 'breast', mardana 'manipulation of the breasts', rati 'sexual intercourse', āsana 'erotic position' etc. makes even erotic manuals read like learned treatises! Turning to a sombre vein, terms like vidhave 'widow', nidhana 'death', śivaikya 'death', divangata 'dead' are generally preferred, in writing at least, to their native-language counterparts.

6.3 Indo-Aryanization and Language Attitudes

As pointed out earlier, the large-scale importation of Indo-Aryan features into the Dravidian languages has not gone unopposed, at least in the literary languages. However, the reasons for the opposition, its intensity, and scope vary from language to language.

In Tamil Nadu, it has turned out to be an emotional issue. Even from the dawn of Dravidian history, the attitude of Tamilians to this issue has been different from that of the speakers of other literary languages. It will be recalled (section 5.1) that Tolkāppiam sanctions the borrowing of only those words that fit into the phonological structure of Tamil. This restriction is not met with in the grammars of the other three literary languages, which have traditionally adopted a more liberal attitude to loan words, welcoming tatsamas and tadbhavas alike, often asserting that theaspirated sounds (mahāprāṇa's) are integral to their language, and recognizing the expanded scripts resulting from incorporation of non-native sounds. All of these grammars do, however, recognize the distinction between native and foreign (dēśya and anyadēśya) words and enumerate eleborate rules for deriving the native forms of the loan-words, but they do not proscribe the use of tatsamas. In fact, they even go one step further and implicitly encourage the use of Sanskrit compounds in their criticism of the employment of hybrid compounds (called arisāmāsa in Kannada and vairisamāsa in Telugu). Further, some grammarians even go out of

their way to show the similarity between the native language and Sanskrit by concocting examples of constructions never used in the language. These grammars are also unlike Tolkāppiam in their adoption of the Paninian model and terminology. (Some of these grammars are, in fact, written in Sanskrit.) Thus the difference in the attitude of the Tamilians on the one hand and the speakers of the other literary languages on the other toward the borrowing of Indo-Aryan features has been pronounced and long-standing.

While the Andhras, the Kannadigas, and the Malayalis have welcomed Indo-Aryan borrowings in principle, they have often reacted to the indiscriminate use of these loans. As noted earlier, the Maṇipravāḷa style never found general grass-roots appeal. On the contrary, there arose a number of counter-movements seeking to bring the literary language closer to the common colloquial idiom. In Kannaḍa, for example, there have been at least three such movements. Āndayya (c. 13th century) composed his Kabbigara Kāva using only native words and assimilated loans; the Vīraśaiva protest (12th century) against the tyranny of Brahmanism found its literary counterpart in the vacanas whose language is relatively free from Sanskrit borrowings. In modern times, there has been a sharp reaction against the Sanskritized style of Kuvempu and others, and a call to eschew unfamiliar Sanskrit words. Similar trends are found in Telugu and Malayalam as well, for example, in the work of Somanātha (who advocated jānu Telugu 'pure Telugu') and Ezhuttacan (a contemporary Malayalam poet).

The target of these movements, however, is only the excessive use of Sanskritic elements--unfamiliar words, long compounds, etc.--but the forms already present in common parlance--both assimilated and unassimilated--are accepted as a proper part of their respective languages. Thus their attitude toward Indo-Aryanization is essentially rooted in considerations of general intelligibility rather than any philosophical objection to borrowing per se. The latter attitude, however, has played an increasingly important role in Tamilian glossopolitics in this century.

In his excellent discussion of the purism movement in Tamil, Annamalai (to appear) observes that the movement was fostered by several factors: the desire to assert the identity of Tamils following the Independence movement which stressed national unity, the discovery of the antiquity of Tamil and its self-sufficiency as demonstrated in the ancient Sangam literature, the change of power-structure from the domination of Brahmans to that of non-Brahmans, among others.

The scope of the "pure Tamil" movement was quite comprehensive. It covered all styles of language, and affected the script, the vocabulary, and even proper names. Most of the borrowed words (including assimilated words) were excised and native Tamil words from the early literature revived to take their place. An informal estimate (quoted in Annamalai, op.cit.) is that the Sanskrit words in use in Tamil writing have come down from fifty percent to twenty percent in the last fifty years. However, there do seem to be more Sanskrit words used in spoken Tamil. This situation is similar to early modern Icelandic, but we know where modern Icelandic has gone.

7.0 CONCLUSION

As we conclude this broad survey of the influence of Indo-Aryan languages on Dravidian, it may not be out of place to step back, as it were, and ask, what in the final analysis has been the role of this process of Indo-Aryanization in the history of Dravidian. This is a very difficult question, not only because it can quickly get transformed into an emotional question, but also because the process is an ongoing one. However, trying to steer clear of value-judgements, one may attempt a partial answer.

First of all, there is no denying the fact that the impact of Indo-Aryan has been both profound and widespread. There is no Dravidian language that has not been touched by this influence, nor is there any 'aspect' of a given language. The Dravidian languages have stood to gain from this process in many ways--their vocabularies have been enriched and their style-range widened. They have developed countless new registers and added to their metrical repertoire. It should not be forgotten that many of the Indo-Aryan words were first introduced by way of literary compositions, some of which are among the finest in Dravidian literature. The Indo-Aryan influence came at a time when the Dravidian languages were beginning to assert their separate identities and aided them in this process. It may be recalled that Indo-Aryanization was probably the key factor in the separation of Malayalam from Old Tamil, the former having adopted a more liberal attitude towards the importation of Indo-Aryan features.

Yet, one does not have to be a purist to concede that Indo-Aryanization has not been an unmixed blessing. Among what in some quarters may be regarded as not quite wholesome effects of this process are the following: the creation of an elite variety based on the incorporation of foreign elements; the emergence of a pedantic literary style, based on the same principle; widespread linguistic slavery manifested in the preference for borrowings over exploitation of native resources; the disappearence of scores of perfectly adequate and expressive native terms under the onslought of the borrowings; and so on.

Granting that the tendency to borrow has been excessive and indiscriminate, it may still be maintained that Indo-Aryanization per se need not take all the blame for the phenomena just mentioned. Socially defined variation is an integral feature of language whether borrowed features serve as the relevant variable or not; variety in literary style is also equally inherent in the medium, literature would be of little interest any other way; words do change their meaning, fall into desuetude, and their place is not always taken by borrowed items. The question, then, is, would the Dravidian languages have developed the variety of registers that they now have, without the benefit of Indo-Aryanization? Probably yes; but the more relevant question is, whatever be our language policy for developing technical terms to serve in contemporary registers, must we get rid of the terms that are already in use, just because they did not once belong to our language? The Tamilians say yes, the other Dravidians say no. So each group goes

in its own way, the former resorting to ancient Tamil, the latter to classical Sanskrit, but the two have one thing in common--they are both attempting to blow new life into antique forms, while a relative newcomer to the scene, English, knocks persistently at the door.

NOTES

*I am grateful to Professors Braj Kachru and Hans Hock for commenting on an earlier version of this paper.

[1]The last part of this claim is not strictly correct. Burrow and Bhattacharya (1970:ix) note that in many areas, the Pengos have lost their own language and have become Indo-Aryan speakers. The authors cite the example of the agricultural tract west of Nowrangpur and adjoining the Bastar border, "where the Pengos no longer speak their own language, though they are aware of its one-time existence." Nevertheless, in so far as the true character and identity of a language resides in its morphology and syntax (apparently basic vocabulary is less impregnable than commonly assumed if the Dravidian case is any proof[see section 4.1 below]), Dravidian languages may be said to be still pretty much intact.

[2]For a cautious questioning of this traditionally accepted position, see Hock (1975).

[3]See References at the end of this paper. The items marked by as asterisk were not available to me at the time of writing.

[4]However, the reliability of the method of glottochronology on which Andronov's claim is presumably based, has been called into question; see, e.g., Bergsland and Vogt, 1962; also note Hock 1976.

[5]I am aware that the use of the term 'caste dialects' has been criticized recently, on the basis of the observation that education rather than caste is the relevant variable. However, if we recall that until recently, the term'educated' could be used interchangeably with 'Brahman', much of the ground for objection disappears. Moreover, sociolinguistic stereotypes are not always commensurate with sociological reality.

[6]This list is limited to those borrowings from Indo-Aryan which have either 1)undergone extensions of meanings or far-reaching phonetic changes or both; or 2) penetrated into non-literary languages, whether or not special meanings are found (Emeneau and Burrow, 1962). Preface).

[7]One can cite numerous examples of this tendency. Sekhar (1969: 161) has observed that Malayalam inscriptions of the 11th and 12th centuries contain examples of concord between modifiers and substantives.

This phenomenon seems to have existed in Telugu as well (Sastri, 1969:271). According to Sekhar, Līlātilakam, the 14th century grammar disapproves of such imitations of Sanskrit syntax in Malayalam. The highly respected 13th century grammar of Kannada, Śabdamaṇidarpaṇa contains an "extrapolated" (anukta) sutra in which it is claimed that even intransitive verbs can be passivized in Kannada.

[8]'Conservatism' as applied to Brahman dialects in Dravidian may be confusing. It is true that the Brahman dialects have traditionally been innovative in acquiring non-native phonology, lexis, etc. However, in so far as the borrowed elements do not undergo the processes of nativization that apply generally in their respective languages, the Brahman dialects are relatively more conservative (see Bright and Ramanujan, op.cit.. and the references cited in that work for further discussion).

REFERENCES

ANDRONOV, M. 1964. On the typological similarity of New Indo-Aryan and Dravidian. Indian Linguistics 25. 119-26.

AIYAR, L.V.RAMASWAMI. 1932. Tulu prose texts in two dialects. Bulletin of the School of Oriental Studies 6.897 -931.

AIYAR, L.V. RAMASWAMI. 1934-35. Semantic divergencies in Indo-Aryan loanwords in South Dravidian. Journal of Oriental Research. Madras, 8-9.

AIYAR, L.V. RAMASWAMI. 1973. A south Indian (Malayalam) evaluation of Sanskrit t(d) and ṭ(ḍ). International Journal of Dravidian Linguistics. 2,1,119-26.

ANNAMALAI, E. (to appear). The movement for linguistic purism: the case of Tamil. Language Movements of this Century in India. Mysore: Central Institute of Indian Languages.

BERGSLAND, KNUT and HANS VOGT. 1962. On the validity of glottochronology. Current Anthropology 3. 115 -53.

BHATTACHARYA, SUDHIBHUSAN. 1957. Ollari: a Dravidian speech. Calcutta: Anthropological Survey of India.

BLOCH, JULES. 1954. The grammatical structure of Dravidian languages (English translation by R.G. Harshe). Poona: Deccan College Post-graduate and Research Institute.

BRIGHT, WILLIAM. 1960. A study of caste and dialect in Mysore. Indian Linguistics 21. 45-50.

BRIGHT, WILLIAM and RAMANUJAN, A.K. 1964. Sociolinguistic variation and language change. Proceedings of the Ninth International Congress of Linguists. Cambridge, Mass.

BURROW, T. andBHATTACHARYA, S. 1970. The Pengo language. Oxford: Clarendon Press.

CALDWELL, ROBERT. 1903. A comparative grammar of the Dravidian or South Indian family of languages. 3rd edition. (Reprinted 1974. Delhi: Orient Publications.)

*DONAPPA, T. 1964. Middle Indo-Aryan loanwords in Telug . Waltair: (Unpublished Ph.D. Thesis) Andhra University.

EMENEAU, MURRAY B. 1962:a. Dravidian and Indian linguistics. Berkeley: Center for South Asian Studies, University of California.
EMENEAU, MURRAY B. 1962:b. Bilingualism and structural borrowing. Proceedings of the American Philosophical Society. 106:5 430-42.
EMEMEAU, MURRAY B. and BURROW, T. 1962. Dravidian borrowings from Indo-Aryan. UCPL 26, Berkeley: University of California.
FOWLER, MURRAY. 1954. The segmental phonemes of Sanskritized Tamil. Language 30. 360-7.
GANESHSUNDARAM, P.C. and VAIDYANATHAN, S. 1958. An evaluation of the Sanskrit loanwords in Tamil from the point of view of Nannul. Indian Linguistics. 19. 63-70.
GODA VARMA, K. 1946. Indo-Aryan loan-words in Malayalam. Mavelikkara: Rama Varma and Bros.
GUMPEREZ, JOHN J. and WILSON, R. 1971. Convergence and creolization: a case from the Indo-Aryan/Dravidian border in India. Dell H. Hymes, ed. Pidginization and Creolization of Languages. 151-67.
HOCK, HANS HENRICH. 1975. Substratum influence on (Rig-Vedic) Sanskrit. Studies in the Linguistic Sciences 5:2. 1975. Urbana: Department of Linguistics, University of Illinois.
HOCK, HANS HENRICH. 1976. Review article on R. Anttila: An introduction to historical comparative linguistics. Language 52. 202-220.
KACHRU, BRAJ B. 1978. Toward structuring the form and function of code-mixing: an Indian perspective. Braj B. Kachru and S.N. Sridhar eds. Aspects of Sociolinguistics in South Asia. Special Issue (No. 16) of International Journal of the Sociology of Language, 21-40.
KITTEL, T. 1894. A Kannada-English dictionary. Mangalore.
KUNJUNNI RAJA K. 1972. The influence of Sanskrit on the Dravidian literatures with special reference fo Malayalam. Arabinda Poddar, ed.: Indian Literature. Indian Institute of Advanced Study.
NADKARNI, MANGESH V. 1970. NP embedded structures in Kannada and Konkani. Ph.D. dissertation, University of California, Los Angeles.
NAYAK, H.M. 1967. Kannada: literary and colloquial--a study in two styles. Mysore: Rao and Raghavan.
* PILLAI, ANAVARTAVINAYAKAM S. 1924. The Sanskrit element in the vocabularies of the Dravidian languages. Mark Collins,ed. Dravidic Studies III. Madras: University of Madras, 27-48.
RADHAKRISHNA, B. 1971. Diglossia in Telugu. Proceedings of the First All India Conference of Linguists. 218-26.
RAMANUJAN, A.K. 1969. The structure of Variation: a study in caste dialects. In Milton Singer and Bernard S. Cohn, eds.: Structure and change in Indian society,pp. 461-74. Chicago: Idine Pub. Co.
SASTRI, KORADE MAHADEVA. 1969. Historical grammar or Telugu--with special reference to Old Telugu, c. 200 B.C.-1000 A.D.. Anantapur: Sri Vendateswara University Post-graduate Centre.
SEKHAR, A.C. 1959. Evolution of Malayalam. Poona: Deccan College Post-graduate and Research Institute.
SHANMUGAM PILLAI, M. 1972. Tamil--literary and colloquial. Linguistics 81. 83-91.

SJOBERG, ANDREE F. 1962. Co-existant phonemic systems in Telugu. Word 18. 269-79.
SJOBERG, ANDREE F. and GIDEON SJOBERG. 1956. Problems in glottochronology: culture as a significant variable in lexical change. American Anthropologist 58:2, 296-300.
SREEKANTAIYA, T.N. 1956. Notes on loans and native replacements in Kannada. American Anthropologist 58. 306-8.
SRIDHAR, S.N. 1975. On the Indo-Aryanization of the Kannada lexicon: a study of the Hindi-Urdu loanwords in Kannada. In H. Van Olphen, ed.: Language borrowing. 99-108. Austin, Texas: Department of South Asian Languages and Literatures, University of Texas.
SRIDHAR, S.N. 1978. On the functions of code-mixing in Kannada. In Braj B. Kachru and S.N. Sridhar, eds.: Aspects of sociolinguistics in South Asia. Special issue (No. 16) of International Journal of the Sociology of Language. 103-11.
SRINIVAS, M.N. 1966. Social change in modern India. Berkeley: University of California Press.
SUBRAMONIAM, V.I. and GANESHSUNDARAM, P.C. 1954. Marathi loans in Tamil. Indian Linguistics. 14. 104-23.
UPADHYAYA, U.P. 1971. Effects of bilingualism on Bidar Kannada. Indian Linguistics. 32.2. 132-138.
*VAIDYANATHAN, S. 1971. Indo-Aryan loanwords in Old Tamil. Madras: Rajan Publishers.

Studies in the Linguistic Sciences
Volume 8, Number 1, Spring 1978

HOW A NOUN CLASS SYSTEM MAY BE LOST: EVIDENCE FROM KITUBA (LINGUA FRANCA KIKONGO)*

Susan U. Stucky

This paper discusses the status of the noun class system in a Bantu lingua franca, Kituba, spoken in southern Zaire. Section 1. outlines the general characteristics of Bantu noun class systems. Noun classes are marked morphologically by prefixes on the noun in both the singular and the plural. Typically, these prefixes trigger agreement for a variety of syntactic categories (i.e. verbs, pronouns, and adjectives). The data in section 2. shows that Kituba has lost many of the plural prefix markers and, of those that remain, many appear to be unstable. In addition, one of the plural markers is being generalized. It is argued that the language is moving towards a system in which the singular is unmarked while the plural is marked. This analysis is supported by evidence from agreement patterns and the behavior of loan words. Section 3. compares the Kituba noun system with that of other Bantu lingua francae. It is pointed out that while Kituba shares many patterns with the other trade languages, it has gone further in the elimination of a noun class system. The claim is then that Kituba is moving towards a nonclass system with an unmarked singular and a marked plural. Furthermore, it is concluded that this change is spreading through the lexicon item by item and proceeds through a variety of mechanisms.

1.0. Introduction

This paper presents a discussion of the status of the noun class system in Kituba, a Bantu language which has undergone considerable simplification over the past several hundred years.[1] Today, Kituba is used primarily as a means of communication between various Bantu language speaking peoples in Bas Zaire, in the Kwangu-Kwilu region of Zaire, and in the southwestern part of Congo-Brazzaville. The data in this paper were obtained from Salikoko Mufwene, a native speaker of Kituba as it is spoken in Bandundu in Zaire.[2] Salikoko refers to his language as either Kikongo ya leta or Kikongo ya bulamatari, which are the two names applied to Kituba in the western Kituba speaking area.

While simplification has taken place in several parts of the grammar, as is typical of trade languages such as Kituba, it is the reduction of the noun class system which is the topic of this paper.

The noun classes in Bantu languages are marked by prefixes on the noun stem. They typically fall into singular and plural pairs so that the noun has an overt morphological marker for the singular noun and a separate overt marker for the plural. It will be shown that the Kituba noun class system has reduced the number of singular-plural prefix pairs significantly and, in addition, that the system is moving towards an unmarked singular-marked plural system through the generalization of one of the existing plural markers. Another characteristic of Bantu noun class systems, agreement, will also be shown to have entirely disappeared. Taken together, these two facts do suggest that very little is left of the noun class system. It is concluded that the noun class system in Kituba is barely functional and is in the process of being eliminated entirely.

1.1. Background

Some background information about Bantu noun class systems will be helpful in following the discussion. As mentioned above, the noun in Bantu languages belongs to a noun class which is marked by a prefix (sometimes ∅) on the noun itself. While the number of classes varies from language to language, there are rarely fewer than ten or more than eighteen in any given language (Guthrie 1967:13). Twenty-three proto classes are proposed by Cole (1971), the largest number of classes reconstructed so far. Within Bantu linguistic tradition, the classes are numbered with reference to the proto-classes. This system will be followed here so that the Kituba evidence may be more easily compared to other Bantu languages.

One of the clearest functions of the noun class system is marking the difference between singular and plural. Often there is a pairing of the singular-plural prefixes and when the term gender is applied to a language of this type, it refers to a pair of prefixes, one which marks the singular and one which marks the plural. Certain behavior of the nouns, however, precludes treatment of the entire system as pairs of singular and plural prefixes. Some classes, such as the locatives, the abstract class, and the liquid class, are not markee for either singular or plural. In addition, two singular prefixes may take a single plural prefix. For example, Proto-Bantu Class 10 (henceforth marked Class*10), a plural marker, is often found as the plural of both Class*9 and Class*11 in a given language. Furthermore, a single prefix in a modern language may be a reflex of two prefixes historically. Finally, there is no necessary correlation between what constitutes a singular-plural pair in one language with this distinction in another. Thus, each prefix must be handled separately. What is significant for this paper is that number is overtly marked in the Bantu languages because both singular and plural forms exhibit a prefix which may be clearly identifiable as separable from the noun stem.

A further characteristic of Bantu noun classes is that they govern agreement in a wide variety of constructions. The categories most often exhibiting agreement include 1) the genitive marker, 2) subject and object pronouns (both clitics and independent forms), 3) adjectives, and 4) verbs. In some cases, a single phonetic form of a prefix may govern two types of

agreement. This is generally taken to be evidence that this prefix is not the marker of a singular class but of two separate classes.

A third feature of noun class systems in Bantu languages is the semantic domain assigned to many of the classes. The correlation between a given noun class marker in a Bantu language and a semantic domain is not as consistent as in the instance of number, so caution is recommended in treating the classes as representative of semantic domains. Nouns do not often fall into categories assigned to the prefix.[3]

In this analysis, nouns in Kituba will be evaluated along the two most consistent parameters, number and agreement, with only slight reference to semantic domain.

2.0. The analysis

To perform the analysis of the Kituba noun system, a list of approximately 450 singular nouns was submitted to the consultant who then provided the corresponding plural forms. The original list was extracted from three sources: Eliet (1953), Swift and Zola (1963), and Fehderau (1966).[4] The first two sources provide a vocabulary from the northern Kituba region. It is interesting to note that there was almost no difference between the singular and plural forms given by the consultant and those provided in the three other sources. Thus, western and northern Kituba are fairly cohesive with respect to the basic vocabulary.

2.1.0. The expression of number in Kituba nouns

This section establishes the relationship of the noun class prefixes exhibited in the languages Fehderau (1966) established to be most closely related to Kituba. These languages include the central Kikongo dialects (Kimanyanga, Kinyombe (north), Ladi and Kisikongo). The Kituba nouns were checked against the central Kikongo vocabulary noted by Laman (1936) to determine what changes in class affiliation have taken place. The resulting system is outlined below. Note that because no agreement remains, the only criterion on which to base the existence of a class will be its pairing with a plural prefix.

2.1.1.0 Class 1 mu-

Very few examples of Class 1 nouns remain in Kituba. Earlier changes in the Kikongo languages show that many of the nouns reconstructed to be in Class*1 have as a prefix only a homorganic nasal. Several scenarios have been proposed for the loss of the vowel of Class 1 mu- and the subsequent assimilation of the nasal. (See Bell 1972, for example). None of these seem to account for this phenomenon in Kituba, however. What is important here is that the appearance of many Class 1 nouns in Kituba with a prefix N- is not an innovation in Kituba. There is another class of nouns which take N- as well. They belong to Class*10. In Kikongo, Class 1 and Class 10 nouns were distinguished by their concord. However, since concord has been

lost in Kituba, the two classes of nouns are indistinguishable. The result is a scission in Class 1 such that some nouns originally in Class 1 now appear in Class 10.

Only two nouns out of the entire lexicon elicited behave as though they belong to Class 1. These are the following:

1.	mu-ana	/mwana/	'child'
	ba-ana	/bana/	'children'
2.	mu-ntu		'person'
	ba-ntu		'people'

Class 1 nouns typically take their plurals in Class 2. These nouns show no sign of change in the plural.

Several other mass nouns which cannot have plurals but which do exhibit a mu- sequence word initially are the following:

3.	mwamba	'sauce'
	mutoki	'perspiration'
	mubisu	'rawness'

Since these forms do not govern agreement nor take plurals, there is no evidence that they are analyzed as containing prefixes. Neither is there any direct evidence that they do not. There is a possibility for mutoki 'perspiration' to be analyzable into a prefix and stem since a verb ku-toka 'to sweat' does exist in Kituba.

2.1.1.1. Class 1a ∅-

Class 1a is a class that did not have a reconstructed prefix. It is usually paired with a Class 2 ba- plural. Class 1a is typically restricted to kinship terms. In Kituba, this class is only of historical interest since other pairs of nouns with no singular marking and a ba- prefixed to the sg. stem have been added to the system. Examples include the following:

4.	bakala/babakala	'man, husband'
	nkuluntu/bankuluntu	'older sibling'
	mama/bamama	'mother'
	tata/batata	'father, sir'

2.1.1.2

Class 2 ba- has already been mentioned with reference to Class 1 and 1a, for which it functions as a plural marker. Ba- has become a plural marker for many nouns in Kituba and it is generally prefixed to the entire noun as in the examples in (4) above. This is the form which, it will be argued, is being generalized as a plural marker.

2.1.1.3 Classes 3 mu- and 4 mi-

Class 3 mu- was originally distinguished from Class 1 mu- in Kikongo for two reasons. First, Class 1 mu- takes as its plural Class 2 ba-, while Class 3 mu- takes as its plural Class 4 mi-. Secondly, Class 1 and Class 3 governed different agreement forms, a distinction which can no longer exist in Kituba since all noun class agreement has been lost. In addition, many of the nouns in Class 3 merged with nouns in Class 9 N- when the vowel of the prefix was lost. Thus, some nouns which originally belonged to Class 3 now belong in Class 9 and the rest are indistinguishable from Class 1 nouns except for their plural pairings.

In Kituba, some change is occuring with respect to the plural marker. In Kikongo, the normal plural form for Class 3 mu- would have been Class 4 mi-. Nouns in Class 3 in Kituba taking Class 4 mi- plurals are the following:

5. mulangi/milangi 'bottle'
   mupepe/mipepe 'breeze, wind'
   musapi/misapi 'finger, toe'
   musuni/misuni 'flesh'
   mutindu/mitindu 'variety, sort'
   mutunga/mitunga 'needle, hospital'
   mudinga/midinga 'smoke'
   mungungu/migungu 'voice, throat'
   mukanda/mikanda 'letter'
   munoko/minoko 'mouth'

For a group of nouns in Class 3 there are alternative plural forms. Note that the first alternative plural consists of ba- prefixed to the singular noun stem. The second plural form is made up of ba- plus the original plural marked by mi-. These are as follows:

6. mungamba/mingamba 'day laborer'
   bangamba
   bamingamba

   mutubi/mitubi 'spokesman, speaker'
   batubi
   bamitubi

   muteki/miteki 'seller'
   bateki
   bamiteki

   muyibi/miyibi 'thief'
   bamiyibi

   munganga/minganga 'doctor'
   bamiganga

mundele/mindele bamindele	'white person, European'
mukombi/mikombi bamikombi	'sweeper'
mwinda/minda bamwinda	'light'

Note that the alternative plural forms were all human nouns but one (the last). Since Class 1/2 typically contains human nouns, it might be argued that these forms were derived by the addition of ba- marking humans. Additional support for this analysis comes from the fact that no alternative agreement forms now exist for distinguishing Class 1 mu- from Class 3 mu-. This seems to be a plausible explanation for the ba+stem forms. The generalization of the ba+mi+stem forms could also be a result of the human marking of ba-. However, we will see later that ba- has been generalized to marking the plural for many nouns not of the human class. In light of this general pattern in the language, forms like ba+mi- stem suggest that the prefixes mu- and mi- are no longer separable and that ba- was prefixed to an already existing plural form for which the mi- was no longer separable. An alternative explanation, that these forms were formed by analogy to French loans where ba- was prefixed to French plurals, is less easy to argue for or against since such an analogical process requires that the speaker be aware that the French loans were plural to begin with. A neighboring lingua franca, Lingala, does, however. Bokamba (personal communication) has pointed out that Lingala (a neighboring lingua franca also has such ba-mi- forms. He suggests that historically the Lingala ba-mi plurals were derived by analogy to the Kituba forms since such plurals are found in only Kinshasa Lingala (the area of greatest contact between Kituba and Lingala speakers) and since these forms are most common among Kituba speakers of Lingala.

2.1.1.4 Class 5 di-

Class 5 has been retained from Kikongo with no change. It has the prefix di- in the singular and takes the plural in Class 6 ma-. There are no exceptions in the data. Class 5 often contains body parts and names of food and plants. Here is a sample:

7.	dikutu/makutu	'ear'
	dibele/mabele	'breast'
	diboko/maboko	'arm, hand'
	dikulu/makulu	'foot, leg'
	dideso/madeso	'bean'
	dinkondo/mankondo	'banana'

2.1.1.5 Class 6 ma-

This plural prefix is more versatile than some plural prefixes because it functioned historically in Kikongo (the base language) as the plural of

nouns in three classes: Classes 5, 11, and 14, and also for some mass nouns, particularly those referring to liquids. The Class 5/6 pairing was noted in (7) above. Examples of the Class 11/6 pairing include only one: lupangu/mapangu 'fence.' The other two found in the vocabularies (lukaya 'leaf,' and lwinda 'cockscomb') were not used by the consultant. Only one Class 14 bu- noun was in the list. However, unlike the Kikongo plural which was in Class 6, the plural in Kituba has shifted to a ba- plural which is prefixed to the singular noun.

8. bu-ala (bwala)/babwala 'village'

The liquid ma- nouns are given below:

9. malafu 'wine, liquor'
maza 'water, river'
mafuta 'oil'

A number of mass nouns also take the Class 6 form.

10. makasi 'anger'
madidi 'coldness'
makelele 'noise'

A few nouns which are not in the liquid class also take a ma- plural prefixed to the singular form.

11. masua/bamasua 'boat'
falanga/mafalanga 'franc'
fofolo/mafofolo 'match'
bafofolo
talatala/matalatala 'glass'

The first of these is undoubtedly of Bantu origin. The second is a borrowing from French but has been widely used in many of the Bantu languages in the area. The second two are more difficult to identify and no explanation is offered. Note that one form has an alternative ba-plural. Another item (clearly a borrowing) for which Laman (1936), Swift and Zola (1963), and Eliet (1953) recorded a ma-plural and which has shifted to a ba-plural in Salikoko's speech, is the form papayi/bapapayi 'papaya.'

Thus Class 6, like Class 4, exhibits a certain degree of instability. Some nouns originally taking a plural in Class 6 have shifted to an unmarked singular form and a marked plural (ba-) form. Others have not. The Class 5-6 pairing, on the other hand, appears to be relatively stable.

2.1.1.6 Classes 7 ki- and 8 bi-

Class 7 ki- nouns took their plural in Class 8 bi- in Kikongo. This pairing still holds for Kituba. Approximately twenty nouns out of the list show this alternation. Several examples follow:

12.	kibaka/bibaka	'wall'
	kifu/bifu	'habit, custom'
	kisalu/bisalu	'work, job'
	kisanunu/bisanunu	'comb'

Certain items have been borrowed from Lingala into Kituba. The Lingala reflex of Class 7 is, however e- and not ki-. Class 8 in Lingala remains bi-. The prefixes of such nouns have not been changed on the borrowed nouns although some alternation between o and u is noted. Two such examples were noted in the list and are given below.

13.	elungi/bilungi	'face'	(Ling. elongi)
	ekeku/bikeku	'statue'	(ling. ekeko)

One could perhaps account for these forms by saying that since Salikoko knows Lingala he uses these forms when speaking Kituba. Bokamba (personal communication) has remarked that a u-o alternation is common in Lingala as spoken by Kituba speakers and furthermore that Kituba u often correlates to Lingala o. Thus, such a correlation seems to be common and accounts for the alternation in (13).

2.1.1.7 Class 9 N-

The reflex of Class*9 ni- in Kikongo is a nasal consonant (N-) which assimilates in place of articulation to a following consonant. Various nasal consonants (ñ , ŋ) appear before vowel initial stems as a result of a number of historical changes which need not concern us here. Reflexes of Class *9 are numerous. This class is often cited as the animal class and there are ample examples of these in the list. In Kikongo (as in many other Bantu languages), Class 9 took a plural form in Class 10. However, Class 9 nouns in Kituba now take a plural ba- prefixed to the singular form.

14.	mbulu/bambulu	'jackal'
	mbuluku/bambuluku	'antelope'
	ngombe/bangombe	'cow'
	mfwenge/bamfwenge	'civit-like animal'
	mfinda/bamfinda	'forest'
	ndunda/bandunda	'vegetables/varieties of vegetables'
	nioka/banioka	'snake'
	nuni/banuni	'bird'

In Kikongo, monosyllabic nouns in Class 9 appear with a prefix vowel i- appearing before the nasal. In Kituba these have all but disappeared. Their plural forms are of the same type of those in (14) above.

15.	(i)mbwa/ba(i)mbwa	'dog'
	(i)mpu/bampu	'hat'
	Ø mvu/bamvu	'year'
	Ø nzo/banzo	'house'

Since Kituba nouns do not govern agreement, the Class 1 mu- and Class 3 nouns with reduced prefixes have fallen together with nouns of Class 9. Both sets now take a marked plural form where ba- is prefixed to the singular. Nouns containing a nasal word initial in the singular and a ba- plural marker prefixed to the singular form constitute, statistically speaking, one of the largest groups of nouns in the list.

2.1.1.8 Class 10 N- or zi-

Class 10, which was either N- or zi- in Kikongo depending on the dialect, was, at first, zi- in Kituba. Shepherd's dictionary (1926) listed Class 10 zi- as the plural of Class 9 but by the time Fehderau did his research (1960's) zi- had disappeared entirely and had been replaced by the prefixation of ba- to the singular form. Class 10, then, has entirely disappeared; its pluralization function has been replaced by the ba- prefix.

2.1.1.9 Class 11 lu-

Class 11 lu- has a variety of plurals associated with it. Class 6 ma- has already been mentioned (See section 2.1.1.5). Class 13 tu- was at one time a regular plural for Class 11. However, there are only a very few examples of Class 11 in the data. Regardless of their plural forms in Kikongo, each instance has been replaced by the marked plural form.

16. lusambu/balusambu 'prayer'
lusadisu/balusadisu 'aid'
lutu/balutu 'spoon'

2.1.1.10 Class 13 ka-

Class 13 ka- apparently has disappeared entirely, there being no examples in the list having this form.

2.1.1.11 Class 14 bu-

Class 14 bu- is attested in only two words. The first bwala/babwala 'village' was discussed in (2.1.1.5). Its plural has been reassigned from Class 6 ma- to the marked plural ba- form. Abstract nouns often appear in this class. By virtue of their abstract meaning they do not have plural forms. One such example is buzoba 'stupidity.' It is related semantically and perhaps morphologically to Kituba zoba 'stupid,' another noun of Bantu origin. Thus it may be that the prefix bu- of buzoba 'stupidity' is still analyzable as a prefix. Still, Class 14 as a plural marker has all but disappeared.

2.1.1.12 Class 15 ku-

Class 15 ku- is the infinitive form. Its status as a noun is not known since no examples appear in the sample. In Kikongo, the ku- form can be

used as a noun and has as its plural Class 6 ma-. As a verbal form in Kituba, it appears in a past progressive construction with an auxiliary verb.

2.1.1.13 Classes 16 va-, 17 ku- and 18 mu-

The status of Classes 16-18 is likewise unclear. In Kikongo they had two functions, one to mark a noun as a locative and the other to construct locative phrases consisting of one of these prefixes and a full noun. The status of the locatives in this second function is not made clear in any of the sources. According to Fehderau "...Lingala forms are more current in the case of these prefixes" (Fehderau 1966:55). That is, Lingala na 'in, at, to, on, by' is preferred to Kikongo (Nimanyanga dialect mu 'in,' ku 'at, to,' va- 'on,' kwa 'by.' The choice of na over the locatives would result in the elimination of Classes 16-18. Na in Kituba is an all purpose particle covering comitative, instrumental, conjunctive, and locative constructions. While locatives were not investigated to any degree in Salikoko's speech several examples using na as a locative turned up in the data. Here is an example:

17. familia na mono meme bika na bwala
family of me past stay loc. village
'my family stayed home'

Locative prefixes also appear attached to noun stems, as noted above, to form locative nouns. None of these were obtained in the sample either. It may be concluded, then, that Classes 16-18 are either not very productive in Kituba or have disappeared entirely.

2.1.1.13 Class 19 fi-

This class is a diminutive in Kikongo, having the plural in Class 8 bi-. In Kituba, according to Swift and Zola (1963), all take a ba- prefixed to the singular to form a plural. The few examples cited by Swift and Zola were not known by Salikoko the Kituba consultant so that no statement can be made about the status of this class.

2.2 Conclusions and hypothesis

Several generalizations can be drawn from the examples just discussed. First, a new method of plural formation has emerged. By marking the singular noun with a ba- to form the plural, part of the system has shifted away from a marked singular/marked plural. Into this group of nouns fall all nouns formerly taking a class 10 plural prefix and those nouns in Classes 1 and 3 which already had a reduced form of the prefixes (a nasal C). The remaining plural classes, Classes 4, 6, 8, and 13 all show some degree of instability. That is, some nouns which took one of these class markers for the plural now take ba- exclusively (prefixed to the singular form) or show some degree of alternation between the newer ba-stem form and the original plural class marker. Thus, it appears that the system is in a state

of flux. For some classes, the movement to a morphologically marked plural is complete. For others (4,6,8, 13), it is not. This change would conform to Greenberg's Universal 35 which claims that there are no languages in which the nonsingular would be (morphologically) marked as opposed to the singular (Greenberg 1966:94), since in every case, marking has taken place in the plural and not the singular when such marking has taken place (e.g. plurals for 3, 9, 11, and 14).

Furthermore, the change has apparently taken place not by a single mechanism, but by several. Thus, the shift of most of Class 3 from a Class 4 plural to a double prefix took place for more human nouns than for non-human nouns. This leads one to suppose that the semantics of Class 2 ba- may have some significance for the change. That is, the addition of ba- may have been, at least in some cases, for semantic reasons. Or, alternatively, it may have been by analogy to class 1a where no singular prefix existed. Secondly, the change in Class 9/10 nouns may have been internally motivated as well. In the original system, Class 9 and 10 nouns would not have been distinguished in the singular and the plural after the elimination of concord. Thus, there may have been pressure to mark plurality.[6] Finally, the classes structurally most vulnerable (i.e. 11, 14) have been affected. Thus, while the overriding reason for the change in the noun class system may have been a drive towards a non-class system, or, at least a simplification of it, the mechanisms are diverse.

Two explanations for the current system in Kituba are possible. The first is that the language is moving towards a non-class system where singulars are unmarked and plurals are marked by ba-. Furthermore, if it is correct that the noun class system is in the process of disintegrating, this data would indicate that the change spreads slowly through the lexicon, item by item and not by wholesale generalization of ba- plural marking to a given class.

Finally, since agreement with noun prefixes has been entirely eliminated, the system has no grammatical function other than to mark singular/plural.

The alternative proposal would be to claim that the noun classes are still present and that the correct interpretation of the unmarked singular is that it contains a ∅- prefix for those nouns. The rest of the forms which conform to the Kikongo pairings (Classes 5/6 and 7/8, for example) would then be evidence for the retention of the noun classes. One problem with this alternative explanation is that it does not explain why there should be variation in the plural forms. More importantly, it does not explain why the variation is consistent.

The first proposal, on the other hand, explains the variation in the plurals. The presence of some nouns which have not changed, this hypothesis would claim, is because the process is not yet completed. Further evidence that this proposal is the correct one (i.e. that the noun class system is disappearing in favor of an unmarked singular-marked plural system) comes

from the treatment of loanwords and from the absence of agreement forms.

2.3 Loanwords

By far the largest number of loanwords come from French. French is widely used in Zaire and has been used since the arrival of the Belgians in the 19th century. A large sample (approximately 120 items) is available. In every instance but one, the plural form involves the prefixation of ba- to the singular form. Here is a partial list of forms used by Salikoko:

18.	are/bare	'(bus) stop'	(Fr. arret 'stop')
	banki/babanki	'bank'	(Fr. banque 'bank')
	bidiki/babidiki	'brick'	(Fr. brique 'brick')
	buteye/babuteye	'bottle'	(Fr. bouteille 'bottle')
	dese/badese	'design'	(Fr. dessein 'design')
	kalaka/bakalaka	'clerk'	(Fr. clerque 'clerk')
	lepose/balepose	'response'	(Fr. reponse 'response')
	maladi/bamaladi	'patient'	(Fr. malade 'sick')
	palado/bapalado	'excuse'	(Fr. pardon 'excuse me')
	takisi/batakisi	'taxi'	(Fr. taxi 'taxi')
	velo/bavelo	'bicycle'	(Fr. velo 'bicycle')
	vwatile/bavwatile	'car'	(Fr. voiture 'car')

The exception noted above is the word for 'bread: dimpa/mampa from the French du pain 'some bread.' It is clearly an exception.

Since alternative loanword adaptation strategies are available for Kikongo languages, the choice of this particular one for Kituba lends support to the claim that an unmarked singular-marked plural is developing.[7] In other dialects of Kikongo, the strategy for adapting French loans is different. If the first syllable of the French noun resembles one of the noun class prefixes, then the noun is assigned to that class regardless of whether the noun was singular or plural in French. Thus, the French biscuit 'cookie' has become bisikiti 'cookies' and its singular form is kisikiti 'cookie' where the sequence bi- of biscuit has been interpreted as the plural prefix bi-. In Kituba on the other hand, the plural is babisikiti 'cookies' where the bi- was not interpreted as the Class 8 prefix. Another example is the French loan for canife 'knife' which became in some Kikongo dialects kanifi/tunifi 'knife' where the ka- of canife was reinterpreted as Class 13 ka- and consequently takes its plural in Class 14. In Kituba, the same noun has been borrowed but it turns up as bakanifi 'knives' in the plural. The fact that the preferred strategy for French loans in Kituba is that of unmarked singular-marked plural lends credence to the claim that the noun class system is no longer very functional in Kituba.

2.4 Agreement

Agreement (or concord as it is usually referred to in the Bantu languages) has been completely eliminated from the Kituba system. In Kikongo, each prefix on the noun had a corresponding concordial prefix (or prefixes),

often phonetically similar, which marked agreement on several classes of words, the main ones being verb, adjective, and the genitive marker. All such agreements no longer appear in Kituba. The following examples illustrate the complete lack of agreement in Kituba verbs. In Kikongo, on the other hand, a prefix would appear on the verb marking agreement with a subject.

19. mwana mene vila — 'the child has been lost'
bana " "

muteki mene vila — 'the seller has been lost'
miteki/ " "
bamiteki " "

ndeke mene tina — 'the bird has flown away'
bandeke " "

diboko mene bukana — 'the arm has broken'
maboko " "

kibulu mene vila — 'the beast has been lost'
bibulu " "

avio mene pela tia — 'the airplane has caught fire'
bavio " " "

bilo mene pela tia — 'the office has caught fire'
babilo " " "

Adjective agreement has also been lost. Normally (in Kikongo) an adjective follows the noun and receives an agreement prefix corresponding to the class of the head noun. In Kituba, adjectives do not follow the noun directly, but are preceded by an associative or genitive marker. Here are some examples:

20. mwana ya fioti — 'small child'
bana " "

muteki ya masonga — 'honest seller'
miteki/" "
bamiteki

ndeke ya kitoko — 'pretty bird'
bandeke " "

diboko ya mbi — 'bad arm'
maboko " "

kibulu ya nene — 'big beast'
bibulu " "

avio ya nene 'big airplane'
bavio " "

bilo ya fioti 'small office'
babilo " "

Note in the examples above that the genitive marker (ya) exhibits no evidence of concord either. In Kikongo, on the other hand, there would have been some agreement morpheme attached to this marker, marking class.

Thus the concordial system, by being fully eradicated, contributes nothing to the retention of noun classes. Furthermore, the one syntactic function assigned to noun classes has entirely disappeared. If the definition of a noun class system includes the process of agreement, then the Kituba system by definition is no longer a noun class system.

2.5 Pronominal forms

The pronominal system does not reflect any difference in noun class either. Instead, there is a tendency towards an inanimate/animate distinction in the third person. Third person animate pronouns are yandi 'he, she, it and bo 'they.' The inanimate pronoun yo serves both the singular and plural functions for inanimates. However, this distinction was one already made in the independent pronominal system of Kikongo and does not reflect an innovation in Kituba.

3. Noun Class Systems in Other Bantu Pidgins

3.1 Reduction of the Noun Class Systems in Other Bantu Pidgins

Previous attempts at summarizing the nature of the reduction of noun classes in Bantu pidgins include Pierre Alexandre's article, "Note sur la réduction du système des classes dans les langues véhiculaires à fonds bantu" (1968), parts of Bernd Heine's book Pidgin Sprachen im Bantu Bereich (1973), and Eyamba Bokamba's article "The impact of multilingualism on language structure: the case of central Africa" (1977).

Alexandre cites several characteristics of the reduction of the system of nominal classes based on an examination of Up-Country Swahili, Lingala and Pidgin A-70. He notes the retention of independent nominal prefixes to show number, the loss of determinatives and concord, with fewer derived prefixes (i.e. mass and abstract markers). However, he also notes a replacing of the grammatical agreement (concord) by a notional one, animate/inanimate. All grammatical agreement in Kituba has been lost. As to the tendency towards an animate/inanimate distinction, the only case found in the pronominal system existed previously in Kikongo. It is not an innovation of Kituba.

Heine notes several other tendancies in data gleaned from Fanagalo, Lingala, several dialects of Swahili, Pidgin A-to, Town Bemba and Kituba.

In addition to the retention of number and loss of concord mentioned above, Heine mentions a tendency to mark the plural while leaving the singular unmarked. In Kituba, this is evident in nouns in the ∅-/ba- classes. The distinction of number is often lost. An example of this is found in the use of absolute pronouns where both 'it' and 'they' are now a single form yawu-yo. Heine also remarks on the loss of class pairs. In Kituba, not only were two plurals lost, but a total of five pairs were eliminated in favor of ∅-/ba-.

From both Alexandre's and Heine's presentation of other Bantu pidgins, it appears that Kituba is now very similar to them, but differs in the extent to which the system has been reduced. Bokamba (1977) in discussing the noun class systems of several dialects of Lingala and Swahili notes several kinds of simplification. In particular he notes for Kinshasa Lingala the prefixation of ba- to already existing plurals. Thus, the language is moving towards a unified plural marking but in a way different from Kituba and all other Bantu lingua francae.

4. Summary

The data in the paper show that the noun classes in Kituba have been severely reduced. It was argued that the language is moving away from a class system towards a system in which the singular is unmarked and the plural marked by an existing plural marker ba- in the language. In addition, it was suggested that the change is proceeding slowly through the lexicon by means of various mechanisms.

NOTES

*This paper represents a continuation of research begun on this topic in 1975 at the University of Kansas. However, the first part of the research was based on secondary sources rather than on informant work. Results of that research were reported on at the 7th African Linguistics Conference (Stucky 1976). I would like to thank Professor Frances Ingemann of the University of Kansas for her careful criticism of that earlier work. Thanks are also due Professors Hans Hock and Charles Kisseberth, and Eyamba Bokamba at the University of Illinois for their help on this more recent research. Any errors are, of course, solely my own responsibility. This research was supported, in part, by a NDFL Title VI Fellowship through the auspices of the African Studies Center at the University of Illinois.

[1]Fehderau, in his dissertation "The Origin and Development of Kituba (Lingua Franca Kikongo)" (1966), tries to demonstrate that Kituba originated as a pidgin arising from a contact situation due to commercial trade and slave traffic beginning in the fifteenth century. Although the impetus for trade may be attributed to the arrival of the Europeans, largely the Portuguese, Fehderau discredits theories pointing to a linguistic basis in Portuguese, English or French (Fehderau 1966:89-92). A further proposal, that

Kituba in its various forms arose by independent parallel development from local languages, is likewise untenable, he maintains. The structural similarities are too numerous for all the dialects of Kituba, and the structure of the local languages is so different from that of Kituba that they must have come from a single source (Fehderau 1966:93). Fehderau then posits as a point of origin, the trading center of Manianga in central Lower Zaire. He maintains that a period of creolization evidenced particularly in urban centers has extended over the last thirty years.

The simplification of the noun class system has, however, taken place within the last thirty or forty years (See Shepherd's 1926 dictionary, for instance). If Kituba is really a pidgin, one wonders why it has taken so long for the noun class system to break down. Furthermore, this breakdown has taken place during Fehderau's period of creolization. Since reduction is not a characteristic of creolization, while expansion is, the evidence from the noun classes mitigates against the classification of Kituba as a pidgin. More internal evidence is needed to support Fehderau's arguments which are based largely on external evidence.

[2]Special thanks are due to Salikoko Mufwene, the consultant. He is a native speaker of Kituba, a rarity, I might add. In addition to Kituba, he speaks Yansi, Lingala, Franch, and English. He reports that he used Kituba in non-academic and non-familial social interaction.

[3]Welmers (1973) lists the following categories as representative of what one finds synchronically in many of the Bantu languages. (The most recent attempt to define the semantics of the Proto-Bantu noun classes is found in Denny and Creider (1976).

1/2	most personal nouns
3/4	trees, plants, other inanimates
5/6	misc., may be augmentative; nouns of style
7/8	misc., may be either augmentative or diminutive; 8 is sometimes used adverbially
9/10	animal names, personal nouns and inanimates
11	attenuative, abstracts
12/13	diminutive
14	abstract nouns
15	infinitive
16-18	locative
19	diminutive

[4]A much larger vocabulary is available in Swartenbroeckx's (1973) Dictionnaire Kikongo et Kituba - Francais. Unfortunately, it is difficult to determine which nouns are exclusively of Kituba origin or in which Kituba speaking area the nouns are used. Another, undoubtedly excellent, source would be Fehderau's Kikongo (ya Leta)-francais-anglais, a dictionary recently published in Zaire (1969). This was not obtainable.

[5]In many Bantu languages, the noun prefixes participate in morphological word formations. Whether these processes are still productive in Kituba is difficult to ascertain. One would suspect that if the prefix system is

no longer marking noun class that it would be unlikely for the prefixes to participate in word formation as well. Given the tremendous difficulties in formulating the notion of productive, no attempt was made to test word-formation in Kituba.

[6]I would like to thank Professor Hock for calling to my attention the possibility of internal motivation for the change in Classes 9/10. In addition, Bokamba has pointed out that Classes 9/10 are generally the first to in most Bantu lingua francae to adept a plural.

[7]A framework for the discussion of the allocation of loanwords within the African class languages can be found in Heine (1968). Heine proposes three different criteria. The first is automatic allocation. All nouns are assigned to a given class because they are loanwords. The second is phonological allocation. If one of the segments is phonologically similar to that of a nominal affix then it is assigned to that class (as in Kikongo). Semantic allocation, assigning a noun to a class because of its common semantic domain, is the third type.

REFERENCES

Alexandre, P. 1967 (1968). "Note sur la réduction des systèmes de classes." Colloque sur la classification nominale dans les langues négro-africaines. Paris: Centre Nationale de la Recherche Scientifique. Colloques internationaux. Sciences Humaines 30:277-290.

Bell, A. 1972. "The development of syllabic nasals in the bantu noun class prefixes MU-, MI-, and MA-." Anthropological Linguistics 14:29-44.

Bokamba, Eyamba. 1977. The impact of multilingualism on language structures: the case of central africa. Anthropological Linguistics 181-202.

Cole, Desmond T. 1971. The history of african linguistics to 1945. In Current Trends in Linguistics, Vol. 7, edited by Thomas A. Seboek. The Hague: Mouton.

Denny, J. Peter, and Chet A. Creider. 1976. The semantics of noun classes in proto-bantu. Studies in African Linguistics 7:1, 1-30.

Éliét, Édouard et. al. 1953. Les langues spontanées dites commerciales du Congo: le monokotuba comparé au lingala et au lari de la région du Pool. Brazzaville: Editions Victor Sammarro.

Fehderau, H.W. 1966. The Origin and Development of Kituba (Lingua Franca Kikongo). Doctoral dissertation (1966), Cornell University, Ithaca, New York.

Fehderau, H.W. 1969. Kikonga (ya Leta). Kinshasa: LECO

Greenberg, Joseph H. 1966. Universals of Language. Cambridge: The M.I.T. Press.

Guthrie, M. 1967. Comparative Bantu: An Introduction to the Comparative Linguistics and Prehistory of the Bantu Languages. Farnborough (Hants), England: Gregg Press Ltd.

Heine, B. 1968. "The allocation of loan-words within the nominal class systems of some togo remnant languages." Journal of African Studies.. 2:130-139.
Heine, B. 1970. Status and Use of African Lingua Francas. Munchen: Welt form Verlag.
Heine, B. 1973. Pidgin Sprachen im Bantu Bereich. Berlin: Dietrich Reimer.
Laman, K. E. 1936. Dictionnaire Kikongo-Francais. Brussels. (Reprinted 1964, Gregg Press).
Nida, Eugene A. and Harold W. Fehderau. 1970. Indigenous pidgins and koines. IJAL 36:146-155.
Polome, E. 1968. "Lubumbashi swahili." Journal of African Languages. 7:14-25.
Stucky, Susan U. 1976. Reduction of the noun class system in kituba (Lingua Franca Kikongo). Present at the Seventh African Linguistics Conference held at the University of Florida, Gainesville, Florida, April 1976.
Swartenbroeckx, Pierre. 1973. Dictionnaire Kikongo et Kituba-Francais. Bandundu: CEEBA.
Swift, L. B. and Zola, E. W. A. 1963. Kituba Basic Course. Washington, D. C. Foreign Service Institute.
Welmers, W. 1973. African Language Structures. Berkeley: University of California Press.

Forthcoming Special Issues

Fall 1978	Vol. 8, No. 2, *Linguistics in the Seventies: Directions and Prospects* (Forum Lectures delivered at the 1978 Linguistic Institute of the Linguistic Society of America) Editor: Braj B. Kachru	$10.00
Fall 1979	Vol. 9, No. 2, *Relational Grammar and Semantics* Editor: Jerry L. Morgan	$5.00
Fall 1980	Vol. 10, No. 2, *Papers on Diachronic Syntax* Editor: Hans Henrich Hock	$5.00
Fall 1981	Vol. 11, No. 2, *Studies in Language Variation: Nonwestern Case Studies* Editor: Braj B. Kachru	$5.00
Fall 1982	Vol. 12, No. 2, *South Asian Linguistics: Syntax and Semantics* Editor: Yamuna Kachru	$5.00

STUDIES IN THE LINGUISTIC SCIENCES

The following issues of SLS are out of print:
Spring 1971, Vol. 1, No. 1, *Papers in General Linguistics;* Fall 1971, Vol. 1, No. 2, *Papers on Hindi Syntax,* Editor: Yamuna Kachru; Spring 1972, Vol. 2, No. 1, *Papers on Syntax and Semantics,* Editor: Georgia M. Green; Fall 1972, Vol. 2, No. 2, *Studies in Baltic Linguistics,* Editors: Hans H. Hock and Michael J. Kenstowicz; Spring 1973, Vol. 3, No. 1, *Papers in General Linguistics;* Fall 1973, Vol. 3, No. 2, *Papers on South Asian Linguistics,* Editor: Braj B. Kachru; and Spring 1976, Vol. 6, No. 1, *Papers in General Linguistics,* Special section: *Topics in relational grammar,* Editor: Jerry L. Morgan.

The following issues are available:

Spring 1974	Vol. 4, No. 1, *Papers in General Linguistics*	$3.50
Fall 1974	Vol. 4, No. 2, *Papers on Phonetics and Phonology* Editors: Charles W. Kisseberth and Chin-W. Kim	$3.50
Spring 1975	Vol. 5, No. 1, *Papers in General Linguistics*	$3.50
Fall 1975	Vol. 5, No. 2, *Papers on Historical Linguistics: Theory and Metho* Editors: Ladislav Zgusta and Hans H. Hock	$3.50
Fall 1976	Vol. 6, No. 2, *Papers on African Linguistics* Editors: Eyamba G. Bokamba and Charles W. Kisseberth	$3.50
Spring 1977	Vol. 7, No. 1, *Papers in General Linguistics*	$3.50
Fall 1977	Vol. 7, No. 2, *Studies in East Asian Linguistics* Editors: Chin-chuan Cheng and Chin-W. Kim	$3.50
Spring 1978	Vol. 8, No. 1, *Papers in General Linguistics*	$5.00

Orders should be sent to:
SLS Subscriptions, Department of Linguistics
4088 Foreign Languages Building
University of Illinois
Urbana, Illinois 61801

Linguistics in the Seventies: Directions and Prospects

Language Change
William S-Y. Wang Paul Kiparsky
Multilingualism
Charles A. Ferguson Roger W. Shuy
Functional Phonology
Morris Halle Jonathan Kaye
Theories of Meaning
John Searle Edward Keenan
Language Universals
James McCawley Bernard Comrie
Language Acquisition
Susan Ervin-Tripp Neilson V. Smith
Collitz Lecture
George Cardona
Linguistic Society of America Lecture
Charles Osgood

Edited by
Braj B. Kachru

Special Issue of
Studies in the Linguistic Sciences
Department of Linguistics
University of Illinois

STUDIES IN THE LINGUISTIC SCIENCES

PUBLICATION OF THE DEPARTMENT OF LINGUISTICS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

EDITORS: Charles W. Kisseberth, Braj B. Kachru, Jerry L. Morgan

REVIEW EDITORS: Robert N. Kantor and Ladislav Zgusta

EDITORIAL BOARD: Eyamba G. Bokamba, Chin-chuan Cheng, Peter Cole, Georgia M. Green, Hans Henrich Hock, Yamuna Kachru, Henry Kahane, Michael J. Kenstowicz, Chin-W. Kim and Howard Maclay.

AIM: *SLS* is intended as a forum for the presentation of the latest original research by the faculty and especially students of the Department of Linguistics, University of Illinois, Urbana-Champaign. Especially invited papers by scholars not associated with the University of Illinois will also be included.

SPECIAL ISSUES: Since its inception *SLS* has been devoting one issue each year to restricted, specialized topics. A complete list of such special issues is given on the back cover. The following special issues are under preparation: Fall 1979: *Relational Grammar and Semantics,* edited by Jerry L. Morgan; *Studies in Arabic Linguistics,* edited by Michael J. Kenstowicz; *Studies in Language Variation: Nonwestern Case Studies,* edited by Braj B. Kachru; *South Asian Linguistics: Syntax and Semantics,* edited by Yamuna Kachru.

BOOKS FOR REVIEW: Review copies of books (in duplicate) may be sent to the Review Editors, Studies in the Linguistic Sciences, Department of Linguistics, University of Illinois, Urbana, Illinois, 61801.

SUBSCRIPTION: There will be two issues during the academic year. Requests for subscriptions should be addressed to *SLS* Subscriptions, Department of Linguistics, 4088 Foreign Languages Building, University of Illinois, Urbana, Illinois, 61801.

Price: $5.00 (per issue)

LINGUISTICS IN THE SEVENTIES: DIRECTIONS AND PROSPECTS

FORUM LECTURES PRESENTED AT THE 1978 LINGUISTIC INSTITUTE OF THE LINGUISTIC SOCIETY OF AMERICA

EDITED BY
Braj B. Kachru

SPECIAL ISSUE OF

STUDIES IN THE LINGUISTIC SCIENCES

VOLUME 8, NUMBER 2
FALL, 1978

DEPARTMENT OF LINGUISTICS, UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS 61801

CONTENTS

PREFACE v

HERMAN COLLITZ LECTURE:

GEORGE CARDONA: Relations between causatives and passives in Indo-Iranian 1

LINGUISTIC SOCIETY OF AMERICA LECTURE:

CHARLES OSGOOD: Conservative words and radical sentences in the semantics of international politics 43

LANGUAGE CHANGE

WILLIAM S-Y. WANG: The three scales of diachrony 63
PAUL KIPARSKY: Analogical change as a problem for linguistic theory 77

MULTILINGUALISM

CHARLES A. FERGUSON: Multilingualism as object of linguistic description 97
ROGER W. SHUY: Multilingualism as a goal of educational policy . . . 107

FUNCTIONAL PHONOLOGY

MORRIS HALLE: Formal vs. functional considerations in phonology. . . 123
JONATHAN KAYE: Functionalism and functional explanations in phonology 135

THEORIES OF MEANING

JOHN SEARLE: Intentionality and the use of language 149
EDWARD KEENAN: On surface form and logical form 163

LANGUAGE UNIVERSALS

JAMES McCAWLEY: Language universals in linguistic argumentation. . . 205
BERNARD COMRIE: Linguistics is about languages 221

LANGUAGE ACQUISITION

SUSAN ERVIN-TRIPP: Whatever happened to communicative competence?. . 237
N. V. SMITH: Lexical representation and the acquisition of phonology 259

PREFACE

The present volume of *Studies in the Linguistic Sciences (SLS)* is devoted to the Forum Lecture Series presented at the 1978 Linguistic Institute of the Linguistic Society of America, hosted by the University of Illinois from June 12 to August 5, 1978.

At the outset, as the director of the 1978 Institute, I felt that the Forum Lecture Series should focus on a few selected topics which would present "the state of the art" in our discipline during the 1970s. I was pleased to find that the Organizing Committee for the Forum Lectures agreed with this proposal. We therefore selected, after long and serious consideration, the following seven areas as the topics for the lectures: language change, multilingualism, functional phonology, theories of meaning, language universals, language acquisition, and neurolinguistics. The idea behind this structure was to provide a synthesis of contemporary approaches to the seven most debated areas in the linguistic sciences.

The papers included here present such an overview. And, as Henry Kahane, Chairman of the Organizing Committee for the Forum Lectures, said while introducing the Series, we wanted to "view each of our seven themes with two sets of eyes, through the medium of two intellects." That explains why we invited two distinguished scholars to discuss each topic. But as we know, in our field — perhaps as in other fields — an overview generally results in a presentation of one's own view. If that subjectivism is present in any of these "state of the art" papers, it is unavoidable, and we did not discourage it.

This volume, however, does not include all of the fourteen Forum Lectures. The two papers on neurolinguistics presented by Harry Whitaker (University of Rochester) and Grace Yeni-Komshian (University of Maryland) could not be included since the authors had other plans for publication.

In addition to the Forum Lectures, this volume includes the Herman Collitz Lecture and the Linguistic Society of America Professor's Lecture. The Collitz Professorship was held by George Cardona (University of Pennsylvania) and the Linguistic Society of America Chair by Charles Osgood (University of Illinois).

There are several agencies and a number of individuals to whom we owe thanks and gratitude for making the Forum Lectures a success, and making it possible to publish them in book form. The very successful and stimulating Lecture Series would not, of course, have been possible without the cooperation of the distinguished scholars who took part. We are indebted to the Forum Lectures Organizing Committee, who planned the Series meticulously, giving due consideration to every subfield of our discipline. The Committee consisted of William O. Dingwall, Georgia M. Green, Braj B. Kachru, Henry Kahane (Chairman), and Michael J. Kenstowicz. Josephine L. Wilcock, a moving spirit behind every undertaking on which our department embarks, took care of every minute detail. In this she received excellent help from Marjorie S. Olson, Secretary of the 1978 Linguistic Institute. Hans Henrich Hock, Assistant Director of the Institute,provided valuable advice for coordinating the lectures with the Institute's heavy academic and social schedule, and in innumerable other ways. William D. Wallace assisted in the editing of the papers. The support of the Division of Applied Linguistics, the Office of International Programs and Studies, and the Department of Linguistics at UI-UC, made it possible to publish this volume.

Braj B. Kachru

Urbana-Champaign

October 1, 1979

RELATIONS BETWEEN CAUSATIVES AND PASSIVES IN INDO-IRANIAN*

George Cardona
University of Pennsylvania

The Sanskrit of Pāṇini's time (ca. 500 B.C.), described in his Aṣṭādhyāyī, has a fully developed system of causative verbs. In Pāṇini's description, causative bases are derived from primitive ones with the affix ṆiC (*i* with two markers), which appears in the form *ay* in certain contexts. For example:

gam 'go' (3 sg. pres. *gacchati*) : *gam-i* (*gamayati*) 'cause to go'
budh (*budhyate*) 'perceive' : *bodh-i* (*bodhayati*)
bhuj (*bhuṅkte*) 'eat' : *bhoj-i* (*bhojayati*)
paṭh (*paṭhati*) 'recite, study' : *pāṭh-i* (*pāṭhayati*)
ās (*āste*) 'be seated' : *ās-i* (*āsayati*)
hṛ (*harati*) 'take, carry' : *hār-i* (*hārayati*)
kṛ (*karoti*) 'do, make' : *kār-i* (*kārayati*)
pac (*pacati*) 'cook' : *pāc-i* (*pācayati*)

Since the Sanskrit Pāṇini describes corresponds closely to the language of Brāhmaṇa and sūtra texts (Cardona 1976a:238), this means the full causative system had developed by late Vedic times. Thieme (1929:17-23) argued that early Vedic lacked true causatives. According to him, the derived bases usually called causatives served originally only as transitives corresponding to intransitives (including passives). Kuryłowicz (1956:88-9, 1964:87) accepted Thieme's thesis in his own attempt to reconstruct the prehistory of this formation. More recently, Jamison (1976) has once more reiterated this position. I shall argue that this generally held view does not withstand scrutiny, hence must be abandoned.

1.0 Let us begin with causatives in the Sanskrit described by Pāṇini.[1]

1.1 The following examples illustrate syntactic features concerning these causatives.

(A)

(1) a. *gacchati mānavako grāmam* 'The boy is going to the village.'
b. *devadatto māṇavakaṃ grāmaṃ gamayati* 'Devadatta has the boy go to the village.'

(2) a. *budhyate putro dharmam* 'The son perceives his duty.'
b. *pitā putraṃ dharmaṃ bodhayati* 'The father makes his son perceive his duty.'

(3) a. *devadatta odanaṃ bhuṅkte* 'Devadatta is eating rice.'
b. *yajñadatto devadattam odanaṃ bhojayati* 'Yajñadatta feeds Devadatta rice.'

(4) a. *śiṣyo vedaṃ paṭhati* 'A student is reciting the Veda.'
b. *adhyāpakaḥ śiṣyaṃ vedaṃ pāṭhayati* 'The teacher is having his student recite the Veda.'

(5) a. *devadattaḥ kaṭa āste* 'Devadatta is seated on a mat.'
b. *devadattaṃ kaṭa āsayati* '. . . has Devadatta sit on a mat.'

(B)

(6) a. *harati bhāraṃ māṇavakaḥ* 'The boy is carrying a load.'
b. *hārayati bhāraṃ māṇavakam* '. . . is having the boy carry a load.'
c. *hārayati bhāraṃ māṇavakena* (idem)

(7) a. *devadattaḥ kaṭaṃ karoti* 'Devadatta is making a mat.'
b. *devadattaṃ kaṭaṃ kārayati* '. . . is having Devadatta make a mat.'
c. *devadattena kaṭaṃ kārayati* (idem)

(C)

(8) a. *odanaṃ pacati devadattaḥ* 'Devadatta is cooking rice.'
odanaṃ pācayati devadattena '. . . is having Devadatta cook rice.'

The non-causal sentence (1a) has the verb *gam* together with the nominative *māṇavakaḥ*[2] and the accusative *grāmam*, respectively denoting the agent and goal of going. The causal sentence (1b) has the causative *gam-i* in construction with *grāmam*, which again refers to the goal of going, and a second accusative: *māṇavakam*, which refers to the agent of going, now the object of causation. Similarly, (2a), (3a), (4a), (5a) have nominative forms that refer to agents of perceiving, eating, reciting, sitting: *putraḥ* 'son', *devadattaḥ* 'Devadatta', *śiṣyaḥ* 'student', *devadattaḥ*. In addition, (2a), (3a), (4a) contain accusative forms used with reference to objects of actions: *dharmam* 'duty', *odanam* 'rice', *vedam* 'Veda'. The verb *ās* of (5a) occurs with the locative *kaṭe* 'mat', denoting the locus where Devadatta is seated. To the nominative forms of all these correspond accusatives in the related causative sentences: (2b) *putram*, (3b) *devadattam*, (4b) *śiṣyam*, (5b) *devadattam*. Note also that (1) - (4) involve verbs of specific semantic groups: movement, perception, consuming, and verbs such as *paṭh*, which take objects that are speech elements, things said or recited.

(6a) and (7a) are like the non-causal sentences of (1) - (4) in that they too contain verbs (*hṛ*, *kṛ*) used with nominative forms referring to agents (*māṇavakaḥ*, *devadattaḥ*) and accusatives referring to objects (*bhāram*, *kaṭam*). Moreover, (6b), (7b) resemble the causative sentences of (1) - (4) in that they too have causative verbs construed with two accusatives. However, (6c), (7c) have no counterparts in (1) - (4). To the nominatives in the non-causal sentences (6a), (7a), correspond here instrumental forms (*māṇavakena*, *devadattena*) instead of accusatives, as in (6b), (7b). The same is true of (8b), which has the instrumental *devadattena* corresponding to the nominative *devadattaḥ* of (8a).

Instrumental forms are also used in passive sentences related to the non-causal sentences (6a) - (8a):

(9) *māṇavakena bhāro hriyate* 'A load is being carried by the boy.'

(10) *devadattena kaṭaḥ kriyate* 'A mat is being made by Devadatta.'

(11) *odanaḥ pacyate devadattena* 'Rice is being cooked by Devadatta.'

The primitive verbs of (1) - (4) also enter into passive constructions, and these too have instrumental forms denoting agents. The passive sentences corresponding to (1a) - (4a) are:

(12) *māṇavakena grāmo gamyate*

(13) *putreṇa dharmo budhyate*

(14) *odano bhujyate devadattena*

(15) *śiṣyeṇa vedaḥ paṭhyate*

in which the singular passive forms *gamyate* (*gam-ya-te*), *budhyate*, *bhujyate*, *paṭhyate* agree in number with the nominative singular forms *grāmaḥ*, *dharmaḥ*, *odanaḥ*, *vedaḥ*.

There are thus two types of causative constructions. One is a double accusative construction: a causative verb occurs with an accusative form referring to the goal, object of the act denoted by the primitive verb and another referring to the agent of this act, the object of causation. In the second construction, an instrumental form is used, as in a passive sentence, with reference to the agent of the act denoted by the primitive verb. The double accusative construction is regular with causative verbs derived from primitive verbs of particular semantic groups, as shown in (1) - (4). The second construction is regular with causatives formed to other transitive verbs. In addition, as shown in (b), two verbs have causatives that are used in either construction. Finally, causatives formed to intransitive verbs such as *ās* follow the pattern shown in (A), with a single accusative referring to the agent of the primitive verb, object of causation.

1.2 I have thus far spoken of transitive and intransitive verbs without specifying criteria whereby given verbs in given contexts may be so classed. There are straightforward criteria for determining in Sanskrit whether or not a verb is transitive, and it will be useful to consider these.

It can be seen from (1) - (4) that the primitive verbs *gam*, *budh*, *bhuj*, *paṭh* can take accusative complements, as can also *hṛ*, *kṛ* of (6), (7) and *pac* of (8). Moreover, these verbs are used in true passive constructions: (9) - (15). Similarly, to

(16) *māṇavako grāmam agamat* 'The boy went to the village.'

(17) *abuddha putro dharmam* 'The son perceived his duty.'

(18) *devadatta odanam abhukta* 'Devadatta ate rice.'

(19) *śiṣyo vedam apaṭhīt* 'A student recited the Veda.'

(20) *māṇavako bhāram ahārṣīt* 'The boy carried a load.'

(21) *devadattaḥ kaṭam akārṣīt* 'Devadatta made a mat.'

(22) *devadatta odanam apākṣīt* 'Devadatta cooked rice.'

with the preterital (third singular aorist) forms *agamat*, *abuddha*, *abhukta*, *apaṭhīt*, *ahārṣīt*, *akārṣīt*, *apākṣīt*, correspond the passive sentences

(23) *agami māṇavakena grāmaḥ*

(24) *abodhi putreṇa dharmaḥ*

(25) *abhojy odano devadattena*

(26) *apāṭhi vedaḥ śiṣyeṇa*

(27) *ahāri bhāro māṇavakena*

(28) *akāri devadattena kaṭaḥ*

(29) *apācy odano devadattena*

These contain the third singular passive aorist forms *agami*, *abodhi*, *abhoji*, *apāṭhi*, *ahāri*, *akāri*, *apāci*, which agree in number with nominative forms denoting objects, just as in (9) - (15). In the same manner, the participial forms *gataḥ* 'gone to', *buddhaḥ* 'perceived', *bhuktaḥ* 'eaten', *paṭhitaḥ* 'recited', *hṛtaḥ* 'carried', *kṛtaḥ* 'made', *pakvaḥ* 'cooked' in

(30) *māṇavakena grāmo gataḥ*

(31) *putreṇa dharmo buddhaḥ*

(32) *odano bhukto devadattena*

(33) *śiṣyeṇa vedaḥ paṭhitaḥ*

(34) *māṇavakena bhāro hṛtaḥ*

(35) *devadattena kaṭaḥ kṛtaḥ*

(36) *odanaḥ pakvo devadattena*

agree in number with nominatives *grāmaḥ* etc. Again in

(37) *devadattena kaṭāv akṛṣātām* 'Two mats were made by Devadatta'.

(38) *devadattena kaṭā akṛṣata* 'Several mats were made by Devadatta'.

the aorist forms *akṛṣātām* (3 du. mid.), *akṛṣata* (3 pl. mid.), used passively, agree in number with the nominatives *kaṭau* (du.), *kaṭāḥ* (pl.), as do the participles *kṛtau*, *kṛtāḥ* of

(39) *devadattena kaṭau kṛtau*

(40) *devadattena kaṭāḥ kṛtāḥ*

Such participles also agree in gender with nominatives referring to objects; *gataḥ* etc. in (30) - (36), (39) - (40) are masculine, as are *grāmaḥ* etc., but *kṛtā* of

(41) *devadattena sthālī kṛtā* 'A pot has been made by Devadatta.'

is feminine, in agreement with *sthālī*.

Let us agree that a verb which occurs in constructions as shown is transitive. A verb such as *ās*, which does not occur in these constructions, is then intransitive. Thus, of itself this verb does not occur with object accusative complements.[3] It does form a passive, as in

(42) *devadattena kaṭa āsyate*

the passive counterpart to (5a), but *āsyate* fails to show number agreement in sentences such as

(43) *puruṣābhyāṃ kaṭa āsyate*

(44) *puruṣaiḥ kaṭa āsyate*

corresponding to

(45) *puruṣau kaṭa āsāte*

(46) *puruṣāḥ kaṭa āsate*

Similarly, to

(47) *devadatta kaṭa āsiṣṭa*

(48) *puruṣau kaṭa āsiṣātām*

(49) *puruṣāḥ kaṭa āsiṣata*

correspond passive sentences

(50) *devadattena kaṭa āsi*

(51) *puruṣābhyāṃ kaṭa āsi*

(52) *puruṣaiḥ kaṭa āsi*

all with the third singular form *āsi*, and

(53) *devadattena kaṭa āsitam*

(54) *puruṣābhyāṃ kaṭa āsitam*

(55) *puruṣaiḥ kaṭa āsitam*

have the neuter singular participial form *āsitam*.

That a verb such as *pac* in (8a), (11), (22), (29), (36) is transitive does not mean active forms like *pacati* must always occur with an object accusative complement or indeed that *pac* must always be transitive. Consider, for example,

(56) *odanaḥ pacyate* 'The rice is cooking.'

(57) *odanena pacyate* [passive equivalent of (56)]

(58) *apācy odanaḥ* 'The rice cooked.'

(59) *apācy odanena* [passive equivalent of (58)]

(60) *odanena pakvam* [alternative to (59)]

A condition for the use of (56) as opposed to (11) is that an agent performing the act of cooking on rice not be mentioned. Moreover, since (57), (59), (60) are formally like (42), (50), (53), one can rightly consider that *pac* of the set (56) - (60) is intransitive.[4] Nevertheless, the fact remains that examples of this type regularly have specific verbs which are otherwise transitive, namely those denoting actions objects in which are produced (e.g. *kṛ* 'make') or undergo a change of state (e.g. *pac* 'cook', *bhid* 'split, break', *lū* 'cut', *muc* 'loose'). Verbs like *ās*, which are regularly intransitive, do not enter into this construction.

Also excluded from constructions of the type (56) - (60) are verbs of movement such as *gam*. These verbs have, in addition, other syntactic features which set them apart. As shown in (1a), the goal of going can be signified by an accusative like *grāmam*. This can also be signified by a dative form such as *grāmāya*:[5]

(61) *gacchati māṇavako grāmāya*

In (30) the participle *gataḥ* is coreferential with *grāmaḥ*, both used with reference to the goal of going. However, *gata* 'gone' can also be agentive, referring to an agent of going, as in

(62) *māṇavako grāmaṃ gataḥ* 'The boy went to the village.'

where *grāmam* denotes a goal.

The verb *smṛ* 'remember, yearn for' behaves in general like *budh*. Thus,

(63) *mātaraṃ smarati devadattaḥ* 'Devadatta remembers his mother with longing.'

has the accusative *mātaraṃ* 'mother', and the passive

(64) *mātā smaryate devadattena*

contains the nominative *mātā*, with which the passive verb *smaryate* agrees in number. However, something one yearns for can also be denoted by a genitive in construction with *smṛ*, as in

(65) *mātuḥ smarati devadattaḥ*

with the genitive *mātuḥ*.[6] Moreover, (65) has a passive counterpart

(66) *mātuḥ smaryate devadattena*

of the type (42): there is no number agreement between *smaryate* and the genitive. If one substituted the dual *svasārau* 'sisters' or the plural *svasāraḥ* for *mātā* in (64), one would also have to substitute the dual and plural verb forms *smaryete*, *smaryante* for the singular *smaryate*, but substituting *svasroḥ* (gen. du.) or *svasṝṇām* (gen. pl.) for *mātuḥ* in (66) entails no change in the verb *smaryate*. Further, since the type (42), in which a passive verb shows no number variation, is typically intransitive, one may rightly consider that *smṛ* is intransitive in (65), (66).[7] This does not alter the fact that *smṛ* in (63), (64) is transitive. Consequently, it is altogether proper to say that the causative *smar-i* of

(67) *mātaraṃ smarayati devadattam*

(68) *mātuḥ smarayati devadattam*

is derived from both an intransitive and a transitive base verb. For (67) and (68) correspond to (63) and (65), respectively.

1.3 The causative constructions spoken of can be reduced to the following schemata, together with related non-causal and passive present constructions:

(69) a.	V N_1-nom. N_2-x[8]	[non-causal]
b.	V-pass. N_1-instr. N_2-nom./x[9]	[passive]
c.	V-*i* N_1-acc. N_2-x N_3[10]	[causative]
(70) a.	V N_1-nom. N_2-acc.	[non-causal]
b.	V-pass. N_1-instr. N_2-nom.	[passive]
c.	V-*i* N_1-instr. N_2-acc. N_3	[causative]

(69c) is the norm for transitive verbs of group (A) and intransitive verbs, (70c) for transitive verbs other than those of (A), and the two constructions are alternatives with causatives formed to the two verbs of (B).

1.4 Though causatives with the affix *i* (*ay*) are freely derived from primitive verbs, it is nevertheless not the case that all bases derived with this affix have only causative value in the Sanskrit of Pāṇini's time. For example, there are non-causative derivates in *-i* such as:

arc-i (*arcayati*) 'praise'
garh-i (*garhayati*) 'revile, deprecate'
chād-i (*chādayati*) 'cover'
dharṣ-i (*dharṣayati*) 'dare'
parc-i (*parcayati*) 'blend'
marṣ-i (*marṣayati*) 'tolerate'
mārj-i (*mārjayati*) 'cleanse'
yoj-i (*yojayati*) 'join'
varj-i (*varjayati*) 'exclude'
vār-i (*vārayati*) 'keep from, ward off from'
śrāth-i (*śrāthayati*) 'loose'
sāh-i (*sāhayati*) 'endure'

It is true that most such bases are not directly cited in the Aṣṭādhyāyī but are known from recensions of the dhātupāṭha (catalog of verbs) which date from centuries after Pāṇini.[11] Nevertheless, it is certain that Pāṇini indeed used such derivates without causative value. His rule 1.4.27: *vāraṇārthānām īpsitaḥ* (*sampradānam* 24) provides that with respect to acts denoted by verbs meaning 'keep from' (*vāraṇārtha*), that which someone wishes to reach (*īpsita*) is classed as sampradāna. The purpose of this classification is to let a rule apply which introduces an ablative ending after a nominal when a sampradāna is to be signified. For example,

(71) *agner māṇavakaṃ vārayati* '. . . keeps the boy from the fire.'

(72) *kūpād andhaṃ vārayati* '. . . keeps the blind man from the well.'

have the ablatives *agneḥ* 'fire', *kūpāt* 'well'. The action noun *vāraṇa* of *vāraṇārtha* in Pāṇini 1.4.27 must be derived from *vār-i*, showing that Pāṇini knew and used such non-causative derived verbs.

1.5 It is important to my subsequent discussion briefly to summarize the general pattern of uses for ātmanepada and parasmaipada affixes--medio-passive and active affixes--in the Sanskrit described by Pāṇini. Passives, whether of the type shown in (9) - (15) or of the type illustrated by (42) - (44), have ātmanepada affixes, and endings such as *te* are regularly preceded by the affix *yaK*.[12] So far as concerns verbs with agentive endings, these are of three general types.

Some verbs take only ātmanepada affixes; for example, the intransitive verb *ās* (*āste*, [5a]) and the transitive verb *bhuj* (*bhuṅkte*, [3a]).[13]

Other verbs take either parasmaipada or ātmanepada affixes, depending on a semantic contrast. If the result of the act signified by the base is intended for the agent, ātmanepada affixes occur; for example, (7a), (8a) are used if Devadatta is making a mat for someone else or cooking rice for someone else to eat; if he were making a mat or cooking rice for himself, one would use *kurute*, *pacate*.[14]

Still other verbs regularly take parasmaipada affixes; for example, the intransitive verb *sad* (*sīdati*) and the transitive verb *paṭh* (*paṭhati*, [4a]).[15]

Causative bases generally follow the second pattern: *kārayati* is used, as in (7b,c), if the causal agent causes something to be made that is not for himself, but *kārayate* is used if this is intended for the causal agent.[16] However, ātmanepada affixes are regularly used in reflexive causatives. For example, corresponding to

(73) *bhṛtyā rājānaṃ paśyanti* 'The king's subjects see him.'

there is a causal sentence

(74) *rājā bhṛtyān darśayate* 'The king shows himself to his subjects.'

in which the causative *darś-i* 'cause to see' is followed by the ātmanepada ending *te* and construed with the accusative *bhṛtyān* 'subjects.'[17]

2.0 With this brief summary of causative and related constructions in the Sanskrit of about 500 B.C. as a background, let us now consider earlier Vedic usage. For the syntax of causatives in Pāṇini's time certainly has a prehistory and reflects changes that occurred in a living language. Moreover, as I have noted, semantic differences play a role in the syntax of causatives in the Sanskrit of Pāṇini's time, so that I shall consider such differences also in connection with causatives of earlier Vedic.

2.1 A goal one reaches is signified by an accusative or a locative in construction with a verb of movement. For example:

(75) RV 1.145.3ab: *tám íd gacchanti juhvàs tám árvatīr víśvāny èkaḥ śṛṇavad vácāṃsi* me 'To him (Agni) go the ladles (i.e. oblations), to him the speedy hymns; he alone will hear all my utterances.'

(76) RV 1.1.4: *ágne yáṃ yajñám . . . viśvátaḥ paribhúr ási/ sá íd devéṣu gacchati* 'Agni, that sacrificial offering which you encompass from all sides goes to the gods.'

In (75), *gacchanti* (3 pl.) takes the accusative complement *tám* 'him', and in (76), *gacchati* (3 sg.) is construed with the locative *devéṣu* 'gods'. Both these complements refer to goals. The causative *gam-i* is also used with an accusative or a locative denoting a goal; for example:

(77) RV 10.145.4cd: *párām evá parāvátaṃ sapátnīṃ gamayāmasi* 'We make the cowife go to the very farthest distance.'

(78) RV 5.5.10c: *tátra havyāni gāmaya* 'Make the oblations go there.'

The accusative *párām . . . parāvátam* in (77) refers to where the wives make a cowife go, the farthest distance, and the locatival *tátra* 'there' of (78) denotes where oblations are supposed to be made to go.

The goal one attains or is made to attain can also be an abstract property, a state, and *gam-i* is used with an accusative referring to such a state. For example, in

(79) TS 2.1.1.1: *sá evaínaṃ bhūtíṃ gamayati* 'He (Vāyu) makes him

(the sacrificer) attain prosperity.'

gamayati occurs with *bhūtim* 'prosperity'. The same is said in TS 2.1.5.5., where it is Indra who causes one to attain prosperity. In a similar context,

(80) MS 2.5.4. (p. 51.15): *yáthaiṣá śríyam aśnutá evám evainaṃ śríyaṃ gamayati* 'As he attains wealth, thus does he cause him to attain wealth.'

gamayati is used with *śríyam* 'wealth'. A parallel passage,

(81) KS 13.6 (p. 188.7-8): *yathaiva sa śriyaṃ prāpnoty evam enaṃ śríyaṃ prāpayati*

has the causative *prāp-i* (*prāpayati*) 'make to attain' with *śríyam*, which is also construed with the non-causative *prāp* 'attain' (*prāpnoti*).

Although in early Vedic verbs of movement such as *gam* usually occur with accusative or locative forms denoting goals, the construction type (61), with a dative referring to a goal, is also used. For example,

(82) AV 16.6.4: *yáṃ dviṣmó yáś ca no dvéṣṭi tásmā enad gamayāmaḥ* 'We make this (bad dream) go to him whom we hate and who hates us.'

has the causative *gamayāmaḥ* construed with the accusative *enad* 'this' and the dative *tásmai* 'to him', the latter referring to that person to whom a bad dream is made to go. This causative construction implies a non-causal sentence in which the goal is also denoted by a dative such as *tásmai*.[18]

2.2 In (75), the subjunctive *śṛṇavat* 'will hear' occurs with the accusative *vácāṃsi* 'utterances, hymns'. A verb of perception such as *śru* 'hear, listen' is also used in construction with a genitive (cf. [65]). In particular, *śru* commonly occurs with a genitive denoting a person to whom one listens. For example,

(83) RV 4.22.10a: *asmā́kam ít sú śṛṇuhi tvám indra* 'Indra, listen well to us.'

has the imperative *śṛṇuhi* construed with the genitive *asmā́kam* 'us'.[19] The verb *cit* 'perceive, know' also is used with either an accusative or a genitive denoting what is perceived. For example,

(84) RV 1.10.2c: *tád índro ártham cetati* 'Then Indra perceives the purpose.'

has the accusative *ártham* 'purpose' with *cetati*, but in

(85) RV 1.2.5ab: *vā́yav índraś ca cetathaḥ sutā́nām . . .* 'Vāyu, you and Indra perceive the Soma juices that have been pressed.'

cetathaḥ (2 du.) occurs with *sutā́nām* (gen. pl.) 'pressed Soma juices'.

Other verbs of perception usually occur with accusative complements. This is true of *dṛś* 'look, see', as in

(86) RV 10.71.4ab: *utá tvaḥ páśyan ná dadarśa vā́cam utá tvaḥ śṛṇván na śṛṇoty enām* 'The one, while looking, has not seen speech, the other, while listening, does not hear her.'

where the perfect *dadarśa* is construed with the accusative *vā́cam* 'speech'. Similarly, *īkṣ* (*īkṣate*) 'look, see' takes an accusative complement, as in

(87) VS 5.34a: *mitrásya mā cákṣuṣekṣadhvam* 'Look at me with the eye of Mitra.'

where the imperative *īkṣadhvam* (2 pl.) occurs with the accusative *mā* 'me'. Causatives are formed to all the verbs mentioned: *śrāv-i* 'make to hear, to listen', *cet-i* 'make to perceive', *īkṣ-i* 'make to see'. In

(88) RV 4.29.3a: *śrāváyéd asya kárṇā vājayádhyai* 'Make his ears hear, so that treasure (booty) be won.'

the imperative *śrāváya* (2 sg.) is used with one accusative: a chanter is told to make Indra's ears (*kárṇā*) hear. What is supposed to be heard is not specified, but this is known from context. The previous verse,

(89) RV 4.29.2ab: *ā́ hí ṣmā yāti náryaś cikitvā́n hūyámānaḥ sotṛ́bhir úpa yajñám* '(Indra,) manly, knowing, comes to the sacrificial offering when he is called by the Soma pressers.'

speaks of Indra coming to the sacrifice when he is called (*hūyámānaḥ* 'being called'), and Indra, characterized as having attentive ears (*āśrutkarṇa*), is elsewhere asked to hear the call of those who call him:

(90) RV 1.10.9a: *āśrutkarṇa śrudhī́ hávam* 'Attentive-eared (Indra,) hear my call.'

The imperfect *ácetayat* (3 sg.) is used with the accusative *acítaḥ* 'who do not perceive, imperceptive' in

(91) RV 7.86.7c: *ácetayad acíto devó aryáḥ* 'The noble god (Varuṇa) made the imperceptive to perceive.'

Gramatically, *acít* is an agent noun that refers to one who does not perceive (*na cetati*). The causative *cet-i* is thus related to the active *cetati*.[20]

In addition, the causative imperatives *darśáya*, *sámīkṣayasva* (2 sg.) 'make to see' are found in double accusative constructions:

(92) AV 4.20.6ab: *darśáya mā yātudhā́nān darśáya yātudhānyàḥ* 'Let me see the warlocks, let me see the witches.'

(93) AV 4.15.3a: *sámīkṣayasva gā́yato nábhāṃsi* 'Let (us) who are chanting see the clouds.'

The accusative plural forms *yātudhā́nān* 'warlocks', *yātudhānyàḥ* 'witches', *nábhāṃsi* 'rain clouds' refer to objects of seeing, and *mā* 'me', *gā́yataḥ* 'those chanting' signify objects of causation, those caused to see these.

2.3 The verb *pā* (3 sg. pres. *píbati*) 'drink' is used with an accusative or a genitive denoting what is drunk. For example,

(94) RV 1.15.1a: *índra sómaṃ píba ṛtúnā* 'Indra, drink the Soma in due time.'

(95) RV 1.4.2b: *sómasya somapāḥ píba* '(Indra,) drinker of Soma, drink the Soma.'

both have the imperative *píba*. Indra is asked to drink Soma juice, and this is referred to by the accusative *sómam* or the genitive *sómasya*. The causative *pāy-i* (2 sg. imper. *pāyáya*) takes two complements: an accusative (e.g. *mádhūni* 'sweet drinks') or a genitive (e.g. gen. sg. *mádhvaḥ* 'sweet drink') denoting what is drunk and an accusative (e.g. *yájatrān* 'those worthy of awe-inspired worship'), as in

(96) RV 3.57.5cd: . . . *ihá víśvāṃ ávase yájatrān ā́ sādaya pāyáyā cā mádhūni* '(Agni,) have all the worshipful ones sit here and drink the sweet drinks.'

(97) RV 1.14.7: *tā́n yájatrāṃ . . . mádhvaḥ sujihva pāyaya* 'Wonderful-tongue (Agni,) have the worshipful ones drink the sweet drink.'

2.4 The verb *vac* (*vivakti*, *vakti*) 'speak, say, utter' occurs with an accusative denoting what is spoken, said, uttered, or said to be and either a dative or an accusative signifying one to whom something is said. For example,

(98) RV 1.74.1b: *mántraṃ vocemāgnáye* 'We would utter a mantra to Agni.'

(99) RV 10.80.7b: *agníṃ mahā́m avocāmā suvṛktím* 'We have uttered a great hymn to Agni.'

In (98), the optative *vocema* occurs with the accusative *mántram* (sacred formula) and the dative *agnáye* 'to Agni', but the aorist *avocāma* of (99) is construed with two accusatives: *suvṛktím* 'hymn' and *agním*.

The causative *vāc-i* is common in Brāhmaṇa and śrautasūtra passages, and this can occur with more than one accusative. For example,

(100) TS 5.4.3.5: *tā́ yájamānaṃ vācayati* '(The Adhvaryu priest) has the patron of the rite say (formulae) to them (the deities).'

has *vācayati* construed with *yájamānam*, referring to the person made to utter formulae, and *tā́ḥ* 'them', which refers to those to whom these are addressed.

2.5 In (96), *sād-i* (*ā́ sādaya* with the preverb *ā́*) 'have . . . sit' occurs with the accusative *yájatrān* and the locatival *ihá*. The base *sad*

'sit', from which *sād-i* derives, is also used in this way, as in

(101) RV 1.188.6c: *uṣā́sāv éhá sīdatām* 'May the two dawns (i.e. Dawn and Night) sit here.'

where *ihá* is used with *ā́* . . . *sīdatām* (3 du. imper.).

2.6 As was shown in 2.2, *śru* and *cit* with active inflexion are used transitively. There are also middle forms of these verbs used intransitively or with passive value. Middle forms of *śru* have the senses 'be heard' and 'be heard of, be famous', as in

(102) RV 1.74.7: *ná yór upabdír áśyvaḥ śṛ́ṇvé ráthasya kácchā́ yád agne yā́si dū́tyàm* 'No hoofbeat sound of horses (as they pull) a moving chariot is ever heard, Agni, when you go on your mission (to the gods).'

(103) RV 8.6.14c: *vṛ́ṣā hy ùgra śṛ́ṇviṣé* 'For, mighty (Indra,) you are famous as a bull.'

where *śṛ́ṇvé* 'is heard' is used with reference to the sound of hoofs (*upabdíḥ*) and *śṛ́ṇviṣé* 'you are heard of (as being . . .), you are famous' is used in connection with Indra. Similarly, *cité* 'is perceived' occurs in

(104) RV 10.143.4ab: *cité tád vāṃ surādhasā rātíḥ sumatír aśvinā* 'Generous Aśvins, your generosity, your benevolence is perceived then (at the time you come to the seat of the ritual).'

Corresponding to such middle forms, *śrāv-i*, *cet-i* are used in the senses 'make to be heard, to be heard of, to be famous' and 'make to be perceived, to appear'. For example,

(105) RV 8.96.12d: *śrāváyā vā́caṃ kuvíd aṅgá védat* '(Chanter,) make heard your speech (so that) perhaps (Indra) will be aware of it.'

(106) RV 7.62.5c: *ā́ no jáne śravayataṃ yuvā́na* 'Young (Mitra and Varuṇa,) make us famous among the people.'

(107) RV 1.3.12ab: *mahó árṇaḥ sárasvatī́ prá cetayati* . . . 'The Sarasvatī makes visible her great flow of water.'

(108) RV 3.34.5c: *ácetayad dhíya imā́ jaritré* '(Indra) made these hymns (thoughts) perceptible to (to be perceived by) the chanter.'

These are of the type (69c), though (105) lacks N_2-x. In (108), moreover, the value of N_2-x is *jaritré* 'to the chanter', a dative. To (108) would correspond a non-causal sentence

(109) *ácitran dhíya imā́ jaritré* 'These hymns became perceptible to the chanter.'

just as

(110) AV 14.1.31c: *bráhmaṇaspate pátim asyaí rocaya* 'Brahmaṇaspati, make (the bride's) husband please her.'

corresponds to a non-causal sentence

(111) *bráhmaṇaspate pátir asyaí rocatām* 'Brahmaṇaspati, may (the bride's) husband please her.'

with a well-known construction: the verb *ruc* (*rocate*) meaning 'be pleasing' occurs with a dative denoting one whom another pleases.[21]

Now *śrāv-i*, *cet-i* in the functions under discussion differ in an obvious way from *sad-i* (2.5), though all three correspond to intransitives. The verb *sad* is regularly intransitive (see 1.2), but *śru* and *cit* can also be transitive, and it is middle forms of these that have intransitive or passive value, while *sad* is regularly inflected actively. The derivate *sād-i* can rightly be considered a transitive counterpart to *sad*. However, *śrāv-i*, *cet-i* cannot be treated so simply without doing injustice. These have two functions. First, they play the role of transitive-causative verbs in relation to middle forms such as *śṛṇve*, *cite*. In addition, they function as causatives to transitives (*śṛṇoti*, *cetati*).

2.7 There are other sets such that an intransitive verb with middle inflexion is paired with a derivate in *-i*, which takes active inflexion. A well-known example is *vṛdh*: *vardhate* 'grows (in strength, age), increases' and *vardhayati* 'makes to grow, to increase'. For example,

(112) RV 6.37.5b: *índro gīrbhír vardhatāṃ vṛddhámahāḥ* 'May Indra, who is of grown might, (still) grow strong through our hymns.'

has the middle imperative *vardhatām*, and

(113) RV 5.11.5cd: *tvā́ṃ giraḥ sindhum ivāvánīr mahī́r ā́ pṛṇanti śávasā vardháyanti ca* '(Agni,) hymns fill you with might, as great streams fill up the Sindhu, and make you grow in strength.'

has *vardháyanti*. Nevertheless, pairs such as *vardhate*, *vardhayati* differ significantly from pairs like *śṛṇve*, *śrāvayati*, considered in the preceding section. To begin with, *śrāvayati* 'makes to be heard' corresponds to a passive *śṛṇve* 'is heard', and the relation between *vardhayati* and *vardhate* is not comparable. At best, one might suppose from comparing (112) and (113) that *índro gīrbhír vardhatām* could mean 'May Indra be made to grow by our hymns', so that *vardhate* could be a quasi-passive to *vardhayati*. However, *vardhate* is an agentive form; compare, for example,

(114) RV 10.161.4ab: *śatáṃ jīva śarádo várdhamānaḥ śatáṃ hemántāñ chatám u vasantā́n* 'Live a hundred autumns, a hundred winters, a hundred springs, increasing in strength.'

which has the participle *várdhamānaḥ* (nom. sg. masc.) 'growing, increasing.' It is, then, proper to say that *gīrbhíḥ* of (112) refers to hymns as causes, not causal agents. In addition, *śrāvayati* 'makes to hear' contrasts with *śṛṇoti* 'hears', but there is not a comparable contrast

between *vardhayati* 'makes to grow' and *vardhati*. On the contrary, these are used in an equivalent manner. For example,

(115) RV 7.99.7c: *várdhantu tvā suṣṭutáyo gíro me* '(Visnu,) let my hymns, of good praise, make you grow strong.'

has the active imperative *várdhantu* used in the same manner as *vardháyanti* in (113).

Similarly, the intransitive verb *ni vartate* (3 pl. imper. *ní vartantām*) 'turn back' contrasts with the transitive verb *ni vartayati* (2 sg. imper. *ní vartaya*) 'make to turn back', but there is no contrast between this and *ni vartati* (3 sg. inj. *ni vártat*):

(116) RV 10.19.3ab: *púnar etā́ ní vartantām asmín puṣyantu gópatau* 'Let these (cows) again turn back, let them thrive in (the possession of) this cow owner.'

(117) RV 10.19.2ab: *púnar enā ni vartaya púnar enā ny ā́ kuru* 'Make these (cows) turn back again, make them be here again.'

(118) RV 1.121.4cd: *yád dha prasárge trikakúm ni vártat ápa druhó mā́nuṣasya dúro vaḥ* 'When the three-humped (bull Indra) turned back (the cows, i.e. waters) in their flow, he opened the doors of the deceiver of man.'

2.8 There are other verbs from which are derived bases in *-i* that are equivalent to active forms of the primitive bases but which differ from *vṛdh*, *vṛt* and such in that their middle forms can have passive value. Consider, for example, *mṛj* 'wipe, clean, polish, curry' (cf. Thieme 1929:19, 22). This verb is transitive by the criteria available in Sanskrit (see 1.2). The active *mārṣṭi* (3 pl. *mṛjánti*) takes an accusative complement denoting an object of cleansing, as in

(119) RV 9.8.4a: *mṛjánti tvā dáśa kṣípaḥ* '(Soma,) the ten fingers cleanse you.'

which has the accusative *tvā* 'you'. The middle *mṛje* (pres. ptcple. *mṛjāná*) is used with passive value, and there is a marked passive *mṛjyate*:

(120) RV 9.109.17: *sá vājy àkṣāḥ sahásraretā adbhír mṛjānó góbhiḥ śrīṇānáḥ* '(Soma,) this steed of a thousand seeds, has rushed, being cleansed by water, blended with cows' milk.'

(121) RV 9.3.3: *eṣá devó vipanyúbhiḥ pávamāna ṛtāyúbhiḥ/ hárir vā́jāya mṛjyate* 'This heavenly (Soma,) as he is filtered, is cleansed by the inspired poets, truth seekers, (like) a bay horse (curried) for treasure (booty).'

In addition, the participle *mṛṣṭá* 'cleansed, curried' is used with reference to an object of cleansing, as in

(122) RV 9.82.2b: *átyo ná mṛṣṭó abhí vā́jam arṣasi* '(Soma,) you rush

towards the treasure prize like a horse that has been curried.'

The base *marj-i* formed to this verb is used in the same manner as the simple active *mārṣṭi*; for example,

(123) RV 9.89.4c: *svásāra īṃ jāmáyo marjayanti* 'The twin sisters (i.e. the fingers of the hand) cleanse him (Soma).'

which in content is fully comparable to (119).

2.9 It is true, as Thieme emphasized, that derived verbs such as *marj-i*, formed to transitive bases other than those considered in 2.1-2.4, are regularly equivalent to simple transitive verbs in the Ṛgveda. It is also true, as has been pointed out (see Delbrück 1888:224-6), that constructions of the types shown in (B), (C) of 1.1 occur in Vedic texts.

The causative *grāh-i* 'cause to seize' occurs frequently. For example,

(124) TS 3.4.8.5: *etád vā asyaí nirṛtigṛhītam/ nirṛtigṛhīta evainaṃ nirṛtyā grāhayati* 'That (part) of this (earth where one practicing a spell against another offers) is seized by Destruction; (the one practicing the spell) causes Destruction to seize him (against whom the spell is intended) at the place seized by Destruction.'

(125) MS 2.5.6 (p. 54.17-8): *prajāpatiḥ prajā́ asṛjata/ tā́ enaṃ sṛṣṭā́ átyamanyanta/ tā atimányamānā váruṇenāgrāhayat* 'Prajāpati created the creatures. When they had been created, they became uppity towards him. He had Varuṇa seize them, being uppity.'

The causative *dāp-i* 'cause to give' occurs in the same construction:

(126) TS 5.4.9.3: *áhnaivā́smai rā́triṃ prá dāpayati rā́triyā́haḥ* '(The Adhvaryu) causes the day to give the night to him (the sacrificer), the night (to give him) the day.'

Similarly,

(127) Ait.Ār. 5.1.5: *rājaputreṇa carma vyādhayanti* 'They have a Kṣatriya pierce the skin.'

has *vyādh-i*, the causative of *vyadh* (*vidhyati*) 'pierce', in the context of part of the Vājapeya ceremony, when Kṣatriyas shoot arrows at a skin. These are instances of construction (70c), with instrumental forms signifying agents of non-causal acts: *nirṛtyā* 'Nirṛti, Destruction', *váruṇena* 'Varuṇa', *áhnā* 'day', *rā́triyā* 'night', *rājaputreṇa* 'king's son, Kṣatriya'.

Causatives in a double accusative construction of type (69c) also occur:

(128) Ait.B 3.46.2: *tad dhaitad eva jagdhaṃ yad āśaṃsamānam ārtvijyaṃ kārayata uta vā me dadyāt uta vā mā vṛṇīteti* 'That is (called) *jagdha* ('eaten', i.e. similar to leftovers) when (the patron of a

sacrifice) has carry out the office of officiating priest (a Brāhmaṇa) who wishes, "May he give me (some money) or select me".'

(129) ŚĀr. 5.8, KBU 3.3.8: *eṣa hy eva sādhu karma kārayati taṃ yam ebhyo lokebhya unnīṣata eṣa u evāsādhu karma kārayati taṃ yam adho nīṣate* 'For this one alone causes him to do good whom he wishes to lead up to those worlds; this one alone causes him to do bad whom he wishes to lead down.'

The accusatives *ārtvijyam* 'office of Rtvij', *sādhu karma* 'good work, deed (sanctioned in traditional lore)', *asādhu karma* 'bad work, deed (forbidden in traditional lore)' refer to objects of the non-causal act denoted by the primitive verb *kṛ* 'make, do', as in Ait.B 5.34.3: *ārtvijyaṃ karoti* 'carries out the office of Ṛtvij'. The accusatives *āśaṃsamānam* 'who wishes', *taṃ yam* 'him whom' refer to objects of causation, signified by the derivate *kār-i* (*kārayate, kārayati*).

Such causatives can also be used in elliptical sentences, where a term of a full construction is to be supplied. For example, ŚB 1.7.1.18: *áthottamā́ṃ dohayitvā* 'Then, after having the last (cow) milked . . .', has one accusative, *uttamā́m* 'last (cow)', but lacks a term referring to the agent of milking. This is known from the sacrificial context; the Adhvaryu, his speech held in check (*vāgyataḥ*), has a non-Śūdra (*aśūdreṇa*) milk (*dohayati*) cows:

(130) KŚS 4.2.22: *vāgyato dohayaty aśūdreṇa*

(130) is itself elliptical, since the objects of milking are not here named and this does not specify the causal agent.

In

(131) a. AV 3.20.8c: *utā́ditsantaṃ dāpayatu* . . . 'And let him cause to give one who does not wish to give.'
b. VS 9.24c: *áditsantaṃ dāpayati* . . . 'He causes to give one who does not wish to give.'
c. TS 1.7.10.1: *áditsantaṃ dāpayatu* . . .

dāp-i occurs with one complement, the accusative *áditsantam* 'who does not wish to give'. Of course, *dā* 'give' is regularly a transitive verb, used with accusative complements denoting things given. But the non-causal sentence corresponding to (131a) would have *ditsati* used absolutely: *áditsantaṃ dāpayatu* = *yó ná ditsati táṃ dāpayatu*. It is understandable, then, that *dāp-i* of (131) occurs in the construction type (69c), with N_1-acc. corresponding to N_1-nom. of the non-causative sentence. The verb *ji* too can be used absolutely: *jayati, vi jayate* 'is victorious, wins'. But *ji* also can be used with an accusative denoting that over which one is victorious, which one conquers, wins. And *jāp-i*, the causative to this verb, is found in a double accusative construction:

(132) VS 9.11: *bṛ́haspate vā́jaṃ jaya bṛ́haspátaye vā́caṃ vadata bṛ́haspátiṃ vā́jaṃ jāpayata/ índra vā́jaṃ jayéndrāya vā́caṃ vadaténdraṃ vā́jaṃ jāpayata* 'Bṛhaspati, win the prize; (drums,) make your sound

to Bṛhaspati, make Bṛhaspati win the prize. Indra, win the prize; (drums,) make your sound to Indra, make Indra win the prize.'

(133) TS 1.7.8.1: *índrāya vā́caṃ vadaténdraṃ vā́jaṃ jāpayaténdro vā́jam ajayit* 'Make your sound to Indra, make Indra win the prize. Indra has won the prize.'

(134) MS 1.11.3 (p. 163.9-10): *índrāya vā́caṃ vadaténdrāya vā́caṃ sáṃ vadaténdraṃ vā́jaṃ jāpayaténdra vā́jaṃ jayeyam* 'Make your sound to Indra, make your sound together to Indra, make Indra win the prize. Indra, I would win the prize.'

2.10 As I noted in 2.6 *áṛṇve* can be used as a passive ('is heard') or in a stative sense ('is heard of, is famous'). Consider now

(135) TS 2.1.4.6-7: *yáḥ pāpmánā gṛhītáḥ syā́t sá āgneyáṃ kṛṣṇágrīvam á labhetaindrám ṛṣabhám/ agnír evāsya svéna bhāgadhéyenopasṛ́ṣṭaḥ pāpmā́nam ápi dahaty aindréṇendriyám ātmán dhatte/ múcyate pāpmáno/ bhávaty evá* 'One who should be seized by evil should immolate a black-necked animal offered to Agni, a bull offered to Indra. Agni, approached with his share (of the offering), burns his evil. Through the (bull) dedicated to Indra, he puts power in himself. He gets loose from evil. Indeed he prospers.'

(136) TS 2.3.13.2: *yáḥ pāpmánā gṛhītáḥ syā́t tásmā etā́m aindrāvaruṇī́m payasyā́m nír vapet/ índra evāsminn indriyáṃ dadhāti/ váruṇa enaṃ varuṇapā́śān muñcati* 'For one who be seized by evil, one should set down this offering of curdled milk dedicated to Indra and Varuṇa. Indra puts in him power. Varuṇa looses him from the noose of Varuṇa.'

The verb of *múcyate pāpmánaḥ* 'He gets loose from evil' in (135) is comparable to *mṛ́jyate* of (121) in that both have the affix *-ya-*. However, in (121) the passive is accompanied by an agentive instrumental, which is absent in *múcyate pāpmánaḥ*. Moreover, the affix *-ya-* bears the high pitch in *mṛ́jyáte*, while in *múcyate* it is the verbal base that is so accented. This form thus has the accent pattern of *jā́yate* 'is born' and statives such as *púṣyati* 'thrives'. It is licit, then, to consider that *múcyate* in (135) is not a true passive, as I have indicated with my translation. This intransitive is related to the transitive *muñcati* 'looses', as in (136). Now,

(137) TS 3.4.1.1: *ví vā́ etásya yajñá ṛdhyate yásya havír ati ríyate* 'His sacrifice fails to prosper the oblation of which is excessive.'

has *ati ríyate* 'is excessive'. This is accented in the same manner as *múcyate*. Moreover, its meaning precludes treating it merely as a passive of *ric* 'leave behind'. This too is a stative, like *ṛdhyate* 'prospers', *ví . . . ṛdhyate* 'fails to prosper'. To the intransitive *ati ríyate* corresponds a transitive *ati recayati* 'makes excessive' (3 sg. opt. *ati recayet*), as in

(138) TS 3.4.1.3: *yád ava dyéd áti tád recayet* 'If he were to cut into (a pregnant animal), then he would make (the offering) excessive.'

This is of the type *śrāv-i* 'make to be heard, to be famous'. Thus, in the language of Brāhmaṇa texts also there are found derivates in *-i* of this kind.

As one would expect from what was pointed out in 1.4, there are also found here derived verbs such as *mārj-i* 'cleanse', without causative value. For example,

(139) ŚB 1.8.1.43: *átha pavítrayor mārjayante* 'Then they wash (with water) in two filters.'

has *mārjayante*, and this is not causative: participants in the rite are said to wash themselves, not to have others wash them. In his commentary on this passage, Sāyaṇa notes that *mārj-i* is here a verb of the tenth class (cf. 1.4 with n 11) and glosses *śodhayanty ātmānam* 'cleanse themselves'.

2.11 There is evidence to indicate that in early Vedic there was, in addition to the morphologically derived causative discussed thus far, a periphrastic causative construction in which a dative of an action noun was used with *kṛ* or *dhā* 'make'. For example,

(140) RV 1.112.8b: *prāndhám śroṇáṃ cákṣasa étave kṛthaḥ* '. . . (Aśvins,) you make the blind to see, the lame to go.'

(141) RV 3.31.19cd: *dúro vi yāhi bahulā́ ádevīḥ svàś ca no maghavan tsātáye dhāḥ* 'Generous (Indra,) avoid the numerous godless deceivers, and make us to win the sun.'

(142) RV 7.79.5c: *vy ucchántī naḥ sanáye dhíyo dhāḥ* 'As you shine, (Dawn,) make our hymns to win.'

However, no periphrastic causative construction was generalized in Indo-Aryan. Only morphologically derived causatives remained the norm,[22] so that these alone will concern us.

2.12 It will be useful now to summarize what I have sketched out thus far and to make an additional point before considering how to account for historical developments. As shown, we have to do with the following kinds of relations.

(a) A transitive verb is related to a derived causative in early Vedic (2.1-2.4):

gacchati 'goes'	*gamayati* 'causes to go'
śṛṇoti 'hears, listens'	*śrāvayati* 'makes to hear, listen'
cetati 'perceives'	*cetayati* 'makes to perceive'
paśyati 'sees'	*darśayati* 'makes to see' [suppletion]
īkṣate 'sees'	*īkṣayate* 'makes to see'
pibati 'drinks'	*pāyayati* 'has . . . drink'
vivakti, bravīti, āha 'says'	*vācayati* 'has . . . say' [suppletion]

(b) An intransitive verb is related to a transitive-causative in

-i (2.5):

sīdati 'sits'	*sādayati* 'seats, sets'

(c) A middle intransitive/passive is related to a transitive-causative in *-i* (2.6 cf. 2.10):

śr̥ṇve 'is heard, is famous'	*śrāvayati* 'make to be heard, famous'
cite 'is perceived, perceptible'	*cetayati* 'makes to be perceived, perceptible'

(d) A middle intransitive is related to a transitive active, which is equivalent to a derivate in *-i* (2.7 cf. 2.10):

vardhate 'grows'	*vardhati, vardhayati* 'makes to grow'
ni vartate 'turns back'	*ni vartati, ni vartayati* 'makes turn back'

(e) A passive, either unmarked, with medio-passive endings alone, or marked, with *ya* before these endings, is related to a transitive active and a derivate in *-i* which are equivalent (2.8 cf. 2.10):

mr̥je, mr̥jyate	*mārṣṭi, marjayati*

(f) In later Vedic, a transitive verb has a passive and a causative in *-i* related to the transitive (2.9):

gr̥hyate 'is seized'	*gr̥hṇāti* 'seizes'	*grāhayati* 'has ... seize'
dīyate 'is given'	*dadāti* 'gives'	*dāpayati* 'has ... give'
vidhyate 'is pierced'	*vidhyati* 'pierces'	*vyādhayati* 'has ... pierce'
kriyate 'is made, done'	*kr̥ṇoti* 'makes, does'	*kārayati* 'has ... make, do'
duhyate 'is milked'	*dogdhi* 'milks'	*dohayati* 'has ... milk'
[*dugdhe* 'gives milk,' n 4]		

The verbs of (a) are transitive in the Sanskrit of Pāṇini's time (1.2). Verbs of the semantic groups in question were used transitively earlier too.

As shown, *śr̥ṇvé* in (102) is passive. The participle *śrutá* 'heard' also has passive value, referring to an object heard:

(143) RV 6.49.3d: *mánma śrutáṃ nakṣata r̥cyámāne* '(Night and Day) come to the devotional hymn, heard (by them) as they are celebrated in verse.'

Though *śrutá* occurs also in the sense 'heard of, famous', this does not alter the fact that it can be passive. The passive *dr̥śyáte* 'is seen' occurs in

(144) AV 7.101: *yát svápne ánnam aśnāmi ná prātár adhi gamyáte/ sárvaṃ tád astu me śivám̐ nahí tád dr̥śyáte dívā* 'What food I eat in a dream is not found in the morning; may that be propitious to me; for it is not seen by day.'

(145) RV 10.88.7ab: *dṛśényo yó mahinā́ sámiddhó rocata diviyonir vibhā́vā* '(Agni,) worthy of being viewed, through his greatness, whose womb is in heaven, who shines afar, has shone when kindled.'

contains *dṛśényaḥ* 'worthy of being seen, viewed'. This is comparable to *uśénya* 'desirable' (RV 7.3.9c) from the verb *vaś* 'desire', which takes object accusative complements. Of the same type is *vīkṣénya* 'worthy of being seen' (RV 9.77.3c). That middle forms of *dṛś* may be considered to mean 'appear' does not change the fact that this verb has passive forms, as does any transitive verb. The passive *pīyáte* 'is drunk' occurs in

(146) AV 5.19.5cd: *kṣīrám yád asyāḥ pīyáte tád vai pitṛ́ṣu kílbiṣam* 'If its milk (the milk of a Brāhmaṇa's cow) is drunk, that is a stain on the Manes.'

(147) RV 1.175.1ab: *mátsy ápāyi te ... matsaró mádaḥ* '(Indra,) get drunk, for you have drunk the intoxicating drink.'

has *ápāyi* 'has been drunk', a passive aorist, and

(148) RV 8.32.16c: *ná sómo apratā́ pape* 'Soma was not drunk without reward.'

contains the passive perfect form *pape*. In addition, the participle *pītá* can refer to an object that has been drunk, as in

(149) RV 6.47.3a: *ayám me pītá úd iyarti vā́cam* 'This (Soma) arouses my speech when drunk (by me).'

It is true that, like English 'drunk', *pītá* can also refer to one who has drunk, but the fact remains that it is used passively. As shown in 2.4 *vac* takes an object accusative complement referring to what one says.

(150) RV 1.77.1ab: *kathā́ dāśemāgnáye kā́smai devájuṣṭocyate bhāmíne gī́ḥ* 'How may we honor Agni? What hymn, pleasing to the god, the shining one, is (to be) uttered?'

has the passive *ucyate* 'is uttered, said', and in

(151) RV 6.34.5b: *índrāya stotrám matíbhir avāci* 'A hymn of praise has been uttered to (this) Indra, through (our) thoughts.'

the passive aorist *avāci* 'has been uttered, said' occurs. Again, in

(152) RV 4.41.1: *índrā kó vāṃ varuṇā sumnám āpa stómo haviṣmā́m .../ yó vāṃ hṛdi krátumām asmád uktáḥ paspárśad indrāvaruṇā námasvān* 'Indra and Varuṇa, which praise with oblation has attained your benevolence? (Which is it,) Indra and Varuṇa, that, full of insight, obeisance, touches your heart when uttered by us?'

uktáḥ 'uttered, said' is used passively. And

(153) RV 6.45.6c: *nṛbhiḥ suvī́ra ucyase* '(Indra,) you are said by

noble men to have a wealth of heroes.'

contains the passive *ucyase* 'you are said ...'

With the above, one may compare similar forms of transitive verbs such as *mṛj* (2.8) or *kṛ*:

(154) RV 5.29.15ab: *índra bráhma kriyámāṇā juṣasva yā́ te śaviṣṭha návyā ákarma* 'Indra most powerful one, take pleasure in the poetic formulae being produced, which we have made, new, for you.'

(155) RV 4.6.11a: *ákāri bráhma samidhāna túbhyam* '(Agni,) you who are being kindled, a poetic formula has been made for you.'

(156) RV 7.61.6d: *kṛtā́ni bráhma jujuṣann imā́ni* 'May these poetic formulae produced (by me) please (you, Mitra and Varuṇa).'

By the criteria which serve to show that a given verb is transitive, then, the above verbs are indeed transitive. Consequently, *śrāvyati* etc. in (a) are to be considered causatives to transitive verbs. That these transitive verbs do have such causatives in early Vedic distinguishes them from others, like *mṛj* (2.8). But then, some of these verbs have an additional property that distinguishes them from other transitive verbs: they can take object complements other than the accusative. This group of items thus forms a fairly well defined subset of transitive verbs.

Verbs of movement such as *gam* also share these features. The verb *gam* occurs with nominals denoting goals, in the accusative, locative or dative, and *gamayati* is related to *gacchati* as its causative. However, in earliest Vedic *gam* and similar verbs had not developed a full passive system. In the Ṛgveda, the participle *gatá* usually refers to an agent of going, and the passive *gamyate* does not occur. Nevertheless, it is also patent that from early times this verb was developing a full passive system. The Ṛgveda itself has an example of *gatá* used with reference to an object, a path one has gone on:

(157) RV 7.58.3: *bṛhád váyo maghávadbhyo dadhāta júyoṣaṇn in marútaḥ suṣṭutī́ṃ naḥ/ gató nádhvā ví tirāti jantúṃ prá ṇaḥ spārhābhir ūtíbhis tireta* '(Maruts,) grant great life force to the generous patrons. Let the Maruts enjoy our praise. It will take a living one across (life) as (does) a path that has been traversed. May you bring us to the fore with your desirable aids.'

Moreover, in

(158) MS 1.6.12 (p. 105.1-2): *jánaṃ bhágo'gacchat tásmād āhur jáno gantavyàs tátra bhágena sáṃ gacchatā íti* 'Bhaga went to the people. Therefore, they say, "One should go to the people; there one comes together with ones share".'

the gerundive *gantavyàḥ* 'to be gone to' is used of a goal to which one should go, the people (*jánaḥ*). Such usage is understandable only if *gam* was indeed treated as a transitive verb. Of course, in the Sanskrit of

Pāṇini's time *gatá* could still refer to an agent or an object, as in (30), (62).

Now, *gatáḥ* in (157) does not refer to a goal. It is used here of a path, road (*ádhvā*) one has taken on the way to a goal. The passive participle *gatá* is thus used here of an object. In this context, it is worthwhile to consider details about the Pāṇinian rule (2.3.12: *gatyarthakarmaṇi dvitīyācaturthyau ceṣṭāyām anadhvani*) which serves to account for pairs of sentences such as (1a) *gacchati māṇavako grāmam*, (61) *gacchati māṇavako grāmāya*. Another rule (2.3.2: *karmaṇi dvitīyā*) says a nominal base is followed by an accusative ending (*dvitīyā*) to signify an object (*karman*). According to 2.3.12, a nominal base is followed by either an accusative or a dative ending (*dvitīyā*, *caturthī*) to signify an object (*karman*) with respect to the act denoted by a verb of movement (*gatyartha*), unless the object is a path, road (*anadhvani*). In his second varttika on 2.3.12 (*āsthitapratiṣedhaś ca*), Kātyāyana remarks that this negation (*pratiṣedha*) applies with regard to a road that one has taken (*āsthita*). That is, the option provided by 2.3.12 applies with reference to a road that is a goal, not a road one takes; the latter is referred to by an accusative, as provided for in 2.3.2. Thus, one says

(159) *adhvānaṃ gacchati* 'He is going on the road, to the road.'

(160) *adhvane gacchati* 'He is going to the road.'

The passive counterpart of (159) is

(161) *adhvā gamyate*

and (157) contains such a passive turn, with the participle *gatá*.

The verb *ṛ* (*ṛchati* 'gets to, reaches') also takes accusative complements. For example, in

(162) TS 5.5.3.2: *yán nyàñcaṃ cinuyāt pṛṣṭhitá enam āhutaya ṛcheyuḥ* 'If one were to set up (the fire) face down, oblations would get to it from the back.'

ṛcheyuḥ (3 pl. opt.) occurs with an accusative (*enam* 'it') referring to the fire into which oblations go. As shown in 2.1, (82) has the causative *gam-i*, used in speaking of making a bad dream go to enemies.

(163) RV 10.164.5de: *pāpó yáṃ dviṣmás tám sá ṛchatu yó no dvéṣṭi tám ṛchatu* 'Let this evil (intent) get to him whom we hate, let it get to him who hates us.'

contains *ṛchatu* 'let ... get to' with an accusative (*tám* 'him') in a similar context. More figuratively, *ṛchati* occurs with an accusative (*etám* 'him') in

(164) TS 2.1.10.2: *ápūtā vā etáṃ vāg ṛchati yám ájaghnivāṃsam abhí śáṃsanti* 'Impure speech gets to him whom people calumniate without his having killed.'

referring to one who is the object of calumny, whom impure speech (*ápūtā vā́k*) gets to. In addition, *ṛcchati* frequently occurs with *śúk* 'pain, affliction' and an accusative denoting that which pain gets to, as in

(165) TS 6.4.1.4: *paśór vā́ ā́labdhasya hṛ́dayaṃ śúg ṛcchati* 'Pain gets to the heart of an immolated animal.'

Compare a passage like

(166) ŚB 3.8.5.8: *paśór ha vā́ ālabhyámānasya hṛ́dayaṃ śúk sam abhy ávaiti* 'Pain gets into the heart of an animal being immolated.'

which has the movement verb *i* 'go' (with preverbs *sam abhy áva*). Further, the past participle *ṛtá* is used with reference to an object to which pain, affliction has gotten. For example,

(167) TS 5.1.4.3: *tásmā́t samā́vat paśūnā́ṃ prajā́yamānānām ā́raṇyā́ḥ paśávaḥ kánīyāṃsaḥ/ śucā́ hy ṛtā́ḥ* 'It is for this reason that, among animals born otherwise the same, wild ones are smaller. For they are afflicted by pain.'

Given *śucā́ ... ṛtā́ḥ* 'gotten to by pain', one must accept that accusatives such as *hṛ́dayam* of *hṛ́dayaṃ śúg ṛcchati* 'pain gets to the heart' in (165) are object accusatives.

The available evidence thus indicates that the primitive verbs of (a), including a verb of movement like *gam*, are transitives. Consequently, *gamayati* etc. of (a) must be considered causatives to transitive verbs. In addition, some of these verbs have passive forms such as *śṛṇve*, and the causatives of (c) are related to these.

Now, by the criteria mentioned, *sad* of (b) is intransitive. Of course, there is no true passive to such a verb, so that the contrast between transitive and passive does not obtain. Moreover, there is here no possible contrast between a transitive and a causative. The same is true of type (d). However, in this case there is also a redundancy in that, despite the absence of an opposition between transitive and causative, one has pairs like *vardhati*, *vardhayati*. In type (e), transitive and passive do contrast, but there is no contrast between transitive and causative. In addition, we have here redundancy both in the passive (*mṛje*, *mṛjyate*) and in the transitive (*mārṣṭi*, *marjayati*). Finally, type (f) is like (a) in that there is here a three-way contrast of passive, transitive, and causative. However, causatives of (f) are demonstrably later than those of (a).

3.0 There are, then, two major questions to be considered: How were causatives of type (f) introduced into the language? How is it that causatives take part in different syntactic constructions? The data sketched above suggest answers to these questions.

3.1 Consider now a development involving passives. From examples I have given, it can be seen that simple middle forms could be used with passive value in early Vedic. Thus: (102) *śṛṇvé* 'is heard', (120) *adbhír*

mr̥jānáḥ 'while being cleansed by water', (148) *pape* 'was drunk'. Similarly,

> (168) RV 10.112.3cd: *asmā́bhir indra sákhibhir huvānáḥ sadhrīcīnó mādayasvā niṣádya* 'Indra, sit down and get yourself drunk with us as you are called by us, friends.'

has *huvānáḥ* 'being called (by us, your friends)', a participial form to *hvé* 'is called', and

> (169) RV 1.92.7b: *diváḥ stave duhitā́ gótamebhiḥ* '(Dawn,) the daughter of heaven is praised by the Gotamas.'

contains *stave* 'is praised'. In addition, there are present and imperfect forms that I call marked passives, in which the affix *ya* precedes middle endings or a participial affix. Thus: (121) *mr̥jyate* 'is cleansed', (89) *hūyámānaḥ* 'being called'. Similarly,

> (170) RV 3.22.1cd: *sahasríṇaṃ vā́jam ... sasanvā́n tsán tstūyase jā́tavedaḥ* '(Agni) Jātavedas, you are praised, being one who was gained treasure in the thousands.'

has the marked passive *stūyase* 'you are praised'. In early Vedic, then, there was a competitive situation, which involved redundancy: either marked or unmarked passives could be used. However, by the period of Into-Aryan represented in Brāhmaṇa texts, the marked passive had all but completely ousted the unmarked passive (see Speyer 1896:49, §167). One is thus practically to the system of the Sanskrit described by Pāṇini, wherein a present or imperfect passive regularly has ātmanepada affixes preceded by *yaK* and third singular aorist forms of the type *akāri* 'has been made', opposed to *akr̥ta* 'has made for himself', are the norm in passive function (see 1.5).

I think it can be considered no accident that, in addition, the stage of Indo-Aryan represented in Brāhmaṇa texts has a fully developed causative system, including causatives of type (f) in 2.12. Not only this, but causative bases have passive forms, as do any transitive verbs. Thus,

> (171) TS 6.5.2.2: *úpopte'nyé gráhāḥ sādyánte'nupopte dhruváḥ* 'Other cups are set on a place where (earth) has been strewn, the Dhruva cup on a place where it has not been strewn.'

> (172) MS 2.3.9 (p. 37.9): *brāhmaṇáḥ pāyayitavyàḥ* 'The Brāhmaṇa is to be given to drink.'

contain *sādyánte* 'are set' and *pāyayitavyàḥ*, a gerundive of *pāy-i* used with reference to one who is to be made to drink. Further,

> (173) ŚB 12.4.1.12: *tád āhuḥ' yásyāgnihotrṁ́ dohyámānā vā́śyeta kíṃ tátra kárma kā́ prāyaścittir íti* 'They say: what is to be done, what is the expiatory procedure, on the part of one whose cow for the Agnihotra ceremony should bleat while one is having it milked?'

has the passive participle *dohyámānā* 'while being caused to be milked', to the base *doh-i*, as in

(174) KŚS 4.14.1: *agnihotrāṃ dohayati* 'He has (someone other than a Śūdra) milk the cow for the Agnihotra' (cf. [130]).

Such passives are also part of the verbal system in the Sanskrit of Pāṇini's time. For example, as (1a) *gacchati māṇavako grāmam* has a passive counterpart (12) *māṇavakena grāmo gamyate*, so also does (1b) *devadatto māṇavakaṃ grāmaṃ gamayati*:

(175) *devadattena māṇavako grāmaṃ gamyate*

3.2 The elimination of unmarked passives in Indic had systematic repercussions in the rest of the verb system. So long as unmarked passive forms of the type *mṛje* coexisted with marked passives of the type *mṛjyate*, the contrast between transitive and passive could be correlated directly with the contrast between active and middle, the affix *ya* being redundant. Once the type *mṛjyate* has become the norm, however, this is not so. Now if a verb has both active and middle endings contrastively, this is correlated only with the distinction between performing an act for ones own benefit and doing this for someone else. Both active and middle endings of the present-imperfect system now can be said to signify agents, as opposed to objective middle endings preceded by *ya*, now required.

In this context, consider again a verb such as *vṛdh*, of type (d) in 2.12. The middle *vardhate* was retained in Sanskrit, as a deponent. On the other hand, the active *vardhati* was gradually eliminated and does not have a place in the Sanskrit of Pāṇini's time. In the emerging system, a transitive *vardhati* could be paired with a middle *vardhate* of the type *kurute* 'makes for himself' or with a true passive *vṛdhyate* of the type *kriyate* 'is made'. Understandably, neither of these was used. Consequently, *vardhati* ceased to be used. In the later system, then, one has only *vardhate*, in which the middle ending signifies an agent of growing, opposed to the causative *vardhayati*. Similarly, the later system has *vartate* and *vartayati*. Note, in addition, that in this way a redundancy was eliminated. In early Vedic, there was no contrast between *vardhati* and *vardhayati* or *ni vartati* and *ni vartayati*. In the emerging system, this lack of contrast is reflected formally, with only *vardhayati*, *ni vartayati*.

3.3 Consider now the consequences of such developments with regard to transitive verbs like *mṛj*, of type (e) in 2.12. As has been shown, in the early system such verbs entered into constructions which are schematically as follows:

(176) a. V- {mid. / *ya*-mid.} N_1-instr. N_2-nom. [passive]
b. V-act. N_1-nom. N_2-acc. [transitive]
c. V-*i* N_1-nom. N_2-acc. [transitive]

With the elimination of the type *mṛje* from the present passive system, we have:

(177) a. V-*ya*-mid. N_1-instr. N_2-nom. [passive]

b. V-act. N_1-nom. N_2-acc. [transitive]

c. V-*i* N_1-nom. N_2-acc. [transitive]

That is, there are now two transitive constructions associated with a single passive construction. Moreover, one of these transitive constructions, (177c), has a verbal base of the type which functions as a causative to a particular subset of transitive verbs, type (a) of 2.12.

I consider that causatives of type (f) had their origin in the transfer of construction (177c) to causative function: (177c) was shifted to (69c) or (70c).

Now, from all that is known about historical developments in language, one must assume that for some time a situation persisted such that, while the verb of (177c) was being used as a new causative, (177c) continued to function also as a transitive construction equivalent to (177b). Indeed, it is difficult to account for the evidence otherwise: in later Vedic, right up to the time of Pāṇini, there are derived verbs such as *mārj-i*, *vār-i*, which do not necessarily function as causatives (see 1.4, 2.10). In this manner, verbs in construction (177c) could still be related to passives in construction (177a), just as an old causative like *śrāvayati* could still be related to a passive; see 2.12(c). Recall now that there is a straightforward relation between sentences in which old causatives of type (a), 2.12, were used and the non-causative sentences connected with them: to a nominative of the non-causative sentence (N_1-nom.) corresponds an accusative (N_1-acc.) in the causative sentence, and other case markings remain the same. By the same relation, a new causative sentence related to (177a) would have N_2-acc. and N_1-instr. In other words, the new causative construction would be (70c).

In the new causative construction, the verb stands as one of three related structures, together with an active and a passive. Hence, there could be additional influence of the old causative construction: the new causative too might enter into a double accusative construction, type (69c). There is evidence to show this occurred at a fairly early time (see 2.9). However, separate constructions were still maintained, and, as things sorted themselves out, by Pāṇini's time only *kār-i* and *hār-i* among the new causatives occur in the double accusative construction, and then only optionally.

Of course, old causatives that were in use from earliest Vedic remained, and in constructions proper to them. In connection with these, recall what I pointed out in 1.2 about (63)-(66). The verb *sṃ̄* can be used transitively. It can also take a genitive complement and occur in a passive construction with such a complement, so that the passive verb shows no variation for number agreement and is of the type *āsyate* in (42). As shown earlier (2.3), verbs of consuming too take accusative and genitive complements in Vedic. And they have passive forms. Thus, *aśyáte* 'is eaten' in

(178) ŚB 2.5.2.23: *pā́treṣu hy áśanam aśyáte* 'For food is eaten in vessels.'

is a passive like *prīyáte* in (146): it agrees in number with *ásanam* 'food'. On the other hand, *aśyate* in

(179) *tásmād etéṣāṃ paśūnāṃ nāśyate* 'Therefore, one does not eat these animals.'

which corresponds to

(180) ŚB 1.2.3.9: *tásmād etéṣāṃ paśūnāṃ nāśítávyam* 'Therefore, one is not to eat these animals.'

is like *smaryate* of (66). It does not agree in number with the genitive plural *etéṣāṃ paśūnām*, just as the neuter singular gerundive *aśítávyam* of (180) shows no gender-number agreement with this genitive. Let us now agree, as I think we must, that sentences such as (1a) and (61) or (94) and (95) are equivalent to each other so far as concerns the relation between the acts in question and the goal or object: in both instances, someone is going to a village and Indra is asked to drink Soma. Further, as shown, verbs of the type (a) in 2.12 are transitive. However, as also shown, the goals, objects of these verbs could be denoted by non-accusative forms. And, as can be seen from (179), (180), such forms could be used also in passive sentences. To such sentences, one could not have a new causative of the type (70c), with an accusative corresponding to the nominative of the passive, for the simple reason that the passive lacked such a nominative. Moreover, as noted in 2.12(a), (c), though causatives such as *śrāvayati* could indeed be related to both an active and a passive, there was a clear semantic difference: *śṛṇóti* 'hears, listens': *śrāvayati* 'causes to hear, listen' as opposed to *śrūyáte* 'is heard': *śrāvayati* 'causes to be heard'. These facts suffice, I think, to account for the fact that construction (70c) was not fully extended, to include also old causatives that entered into the double accusative construction, eliminating (69c).

One should, nevertheless, expect some encroachment of (70c), under specific conditions: if a verb of the semantic groups for which (69c) was the norm were used, possibly with specialized meaning, regularly with an accusative complement. There is evidence to show such extension. For example,

(181) TS 6.4.1.4-5: *sā́ hṛdayaśūlám abhí sám eti/ yat pṛthivyā́ṃ hṛdayaśūlám ud vāsáyet pṛthivīṃ śucā́rpayed yád apsv àpáḥ śucā́rpayet/ śuṣkasya cārdrásya ca saṃdhā́v ud vāsayaty ubháyasya śā́ntyai/ yáṃ dviṣyā́t tám dhyāyet/ śucaívainam arpayati* 'It gets to the heart spike. If (the Adhvaryu) were to remove the heart spike onto the ground, he would cause pain to get to the earth, if onto water, he would cause pain to get to the water. He removes it to a juncture of dry land and wet land, to let both have peace. Let him (then) think of one he hates. (Thereby,) he causes pain to get to him.'

has *pṛthivīṃ śucā́rpayet* 'he would cause pain to get to the earth', in which the causative *arp-i* (3 sg. opt. *arpayet*), seen also in *śucā́ ... enam arpayati* 'he causes pain to get to him', is construed with an accusative (*pṛthivīm*, *enam*) that refers to the object afflicted by pain and an instrumental (*śucā́*) referring to pain, affliction, which is caused to get to

these. As can be seen from *sá* 'it (i.e. pain)', this passage is part of a larger context. It continues immediately after (165), where *hṛ́dayaṃ śúg ṛ́chati* occurs. Patently, *arp-i* is the causative to *ṛ*, a verb of movement. The primitive verb *ṛ́chati* in usages such as (165) regularly takes only an object accusative complement, which implies a nominative in a passive construction together with an instrumental referring to pain which gets to someone (cf. [167]). Accordingly, one finds *arp-i* in construction (70c).

There is also evidence to show that, at least in some areas, the causative construction (70c) was extended to other verbs which originally entered into double accusative causatives. In vārttika 5 on Pāṇini 1.4.52 (*adikhādinīvahīnāṃ pratiṣedhaḥ*), Kātyāyana remarks that the agent of certain acts, among them those denoted by *ad*, *khād* 'eat', should not be classed as object with respect to causatives. In his discussion of this vārttika, Patañjali gives the following examples (Abhyankar 1962:337, lines 15, 18):

(182) a. *atti devadattaḥ* 'Devadatta is eating.'
b. *ādayate devadattena* '... is feeding Devadatta.'

(183) a. *khādati devadattaḥ*
b. *khādayati devadattena*

Further, in vārttika 7 on 1.4.52 (*bhakṣer ahiṃsārthasya*) Kātyāyana notes that the same prohibition should be made for the causative of *bhakṣ* meaning 'eat'. By the third century B.C., then, these three verbs of consuming formed causatives which could occur in construction (70c). And by this time these verbs regularly took only accusative complements.

However, so far as I know, this construction had not been fully generalized by Pāṇini's time. Indeed, verbs of the semantic groups which had construction (69c) in early Indo-Aryan continued to enter into a causative construction different from that of other transitive verbs, and do so to this day (see note 22).

3.4 In sum, the available evidence requires one to accept that earliest Indic had true causatives, formed to particular transitive verbs. New causatives were then developed at a fairly early time, certainly by the times of Brāhmaṇa works, leading to the system known from Pāṇini's era.[23] I suggest the new causative construction is to be accounted for as shown in 3.3.

4.0 At the beginning of this paper, I pointed out that according to the view now prevalent early Vedic did not have true causatives. Let me now turn to the arguments advanced in favor of this position.

4.1 Jamison (1967:126) states the following: 'Though in early Skt. (= RV, AV, early Brāhmaṇas) there existed no specifically causative means of expression, either morphologic or periphrastic, in the later language (= later Brahmanas, Epic, and Classical) causativity could be expressed morphologically by means of a verbal formation with -aya-suffix. The -aya-causative of this later period is the formal continuation of the -áya-verb of the older Vedic period. However, this earlier formation was not a

causative in function but merely had tr. [transitive, GC] value.' Later (1976:127), Jamison remarks that the fundamental property of the formation in question in the later language is causativity, not transitivity. One of the pieces of evidence which leads to this conclusion is that, according to Jamison, non-causative transitive derivates such as *chād-i* in

> (184) RV 6.75.18a: *mármāṇi te vármaṇā chādayāmi* 'I cover your vulnerable areas with armor.'

cannot be formed in the later language. The evidence fails to support Jamison's claims. In view of even the select examples given in 2.9, her statement concerning the absence of causatives in the early language must be considered false, so far as concerns the Atharvaveda and Brāhmaṇas. Moreover, since derivates in *-i* were not made exclusively causative in Sanskrit, right up to Pāṇini's time, her claim about the later language is also factually unacceptable.

Jamison's major claims are two. First, she says derivates in *-i* were formed in the early language only to intransitive verbs, so that the central function of these derived bases was then transitivity. Second, true causatives arose, according to her, through a syntactic analogy due to 'the inherent ambiguity' (1976:133) of accusatives in sentences with two accusative forms. To support her claim that early derivates in *-i* could only be formed as transitives to fundamentally intransitive verbs, Jamison has of course to do something about the verbs of my type (a), 2.12. Her procedure is simple. She merely says (1976:130) that verbs which, though they do take accusative complements, also take others are thereby intransitive, and concludes: 'Thus, according to my definition, when a noun can occur in the AC [accusative case, GC] or in another case in the same meaning, the noun is not a DO [direct object, GC], and the verb with which it appears is not transitive, even at those times when the non-object complement is in the AC.' Further, Jamison strongly emphasizes the type *vardhayati* (see 2.7), which occurs with a single object accusative, and says (1976:126), 'In the early language the defining syntactic feature of the *-áya*-transitive was that it could be construed with a single AC direct object . . .' She also remarks (1976:126) that a verb such as *vardh-i* in RV 1.36.11d: ... *agníṃ vardhayāmasi* 'We make Agni grow strong' had 'another secondary syntactic feature' in that such a sentence could 'be perceived to be functioning as causative to a corresponding intr. sentence built to the same underlying root . . .' According to Jamison, then, a sentence such as AV 12.3.34d: *enam* [var. *etám*] ... *gamayā́ntam* (which she translates [1976:132] 'Make him go to the end') has only one object accusative: *enam* 'him'. The accusative *ántam* is not objective, either here or in the corresponding non-causative sentence *sá gacchaty ántam*, which Jamison treats as intransitive. However, she goes on to say (1976:133), sentences like the last cited have surface structures identical to sentences with transitive verbs, both containing accusatives. And due to the ambiguity of the accusative, a sentence such as *enam ... gamayā́ntam* could then be reinterpreted as containing two object accusatives. In this manner, she concludes (1976:133), '. . . causativity, came to be perceived as the central function of the *-áya*-formation,' although earlier this was merely (1976:128), '. . . an accidental byproduct of the relation between a fundamentally tr. and a fundamentally intr. verb.'

It is crucial to Jamison's thesis that the accusatives in sentences such as (75) *tám ... gacchanti ...*, *víśvāny ékaḥ śṛṇavad vácāṃsi me*, (84) *... índro ártham cetati*, (86) *... tvaḥ ... ná dadarśa vácam*, *... tvaḥ ... na śṛṇoty enām*, (87) *... mā ... tikṣadhvam*, (94) *... sómaṃ pība ...* not be object accusatives. Moreover, she rightly stresses (1976:130) the need to distinguish between object accusatives and others 'without appealing to semantic intuitions,' that is, on formal grounds. However, the single formal criterion she uses is not sufficient. Simply to say that a verb which can take a non-accusative complement in addition to an accusative is thereby intransitive by definition is not justifiable. Recall now that (114) contains the participial form *várdhamānaḥ*, to the intransitive verb *vṛdh* (2.7), the intransitive verb *jīv* (2 sg. imper. *jīva*) 'live', and the accusatives *śatam ... śarádaḥ* 'a hundred autumns', *śatáṃ hemántān* 'a hundred winters', *śatám ... vasantān* 'a hundred springs'. Similarly,

> (185) TB 1.1.3.9: *sò'śvatthé saṃvatsarám atiṣṭhat* 'He (Agni) remained in the Aśvattha tree a year.'

has the intransitive *sthā* (3 sg. impfct. *atiṣṭhat*) 'stand, remain', the locative *aśvatthé*, and the accusative *saṁvatsarám* 'a year.' Again, in

> (186) ŚB 11.5.1.11: *tán ma ékāṃ rā́trim ánte śayitā́se* 'Then you will lie near me one night.'

the intransitive *śī* (2 sg. fut. *śayitā́se*) 'lie, sleep' occurs with the locative phrase *me ... ánte* 'near me' and the accusative phrase *ékāṃ rā́trim* 'one night'. That *jīv*, *sthā*, *śī* occur with accusatives in such examples does not, or course, suffice to make them transitive verbs. The accusatives found here have specific semantic features. In (114), (185), (186), they refer to times associated with the acts in question, during which these take place or are supposed to, and in

> (187) RV 2.16.3cd: *ná te vájram ánv aśnoti kā́ścanā yád āśúbhiḥ pátasi yójanā purú* '(Indra,) no one reaches your vajra when you fly many yojana with your fast (horses).'

the accusative *yójanā purū* 'many yojana (measures of distance)' used with *pat* 'fly' refers to a distance. In effect, these are adverbial accusatives, traditionally called accusatives of extent, and they are neutral in the sense that they can be used equally with intransitive and transitive verbs. Thus, *prathamā́ṃ rā́trim* 'the first night' in

> (188) SB 12.8.2.8: *āśvinéna páyasā prathamā́ṃ rā́trim pári ṣiñcati* 'He sprinkles (the wine) the first night with the milk dedicated to the Aśvins.'

occurs with the transitive verb *pári ṣiñcati* 'sprinkles round'. Similarly,

> (189) RV 1.98.2d: *sá no dívā sá riṣáḥ pātu náktam* 'May he (Agni) protect us from harm day and night.'

has the transitive verb *pā* 'protect' with the object accusative *naḥ* 'us' and the adverbial phrase *dívā ... náktam* 'day and night', in which the

accusative *náktam* is coordinated with the instrumental *dívā*. Such adverbial elements remain unchanged in passive sentences. For example,

(190) RV 10.95.4cd: *ástaṃ nanakṣe yásmiñ cākán dívā náktaṃ śnathithā́ vaitaséna* '(Urvaśī) came to the house (of Purūravas), in whom she took pleasure, pierced by his rod day and night.'

has the passive participle *śnathitā́* 'pierced' with the adverbial phrase *dívā náktam*. The situation is quite different concerning the accusatives which occur with *gam*, *śru*, *cit*, *dṛś*, *√kṣ*, *pā* ('drink') in (75), (84), (86), (87), (94), and similar sentences. To begin with, these accusatives do not fit into any one semantic category; they are not like adverbial accusatives of time and distance. In addition, the verbs with which these accusatives occur are found in passive constructions such that passive verb forms show number and gender agreement with nominative forms corresponding to accusatives in non-passive sentences (see 2.12). Thus, if one considers the criteria which should be considered, one must conclude that these verbs have to be treated as a subset of transitive verbs. By singling out one syntactic feature and omitting others, Jamison fails to take into account the necessary facts. Hence, the crucial part of her thesis fails. Given this, there is no need to go into her explanation concerning how speakers are supposed to have perceived things.

Jamison also suggests an explanation for the use of instrumental forms in the construction type (70c). She says (1976:134, n 3), 'I consider this type of expression a secondary development of the one with double AC, designed to avoid the ambiguities inherent in having both DOs in the same case.' This too I consider unacceptable. After all, despite possible ambiguity, the double accusative causative construction was maintained. Moreover, the distribution of the two causative constructions used in the Sanskrit of Pāṇini's day shows a distinction among transitive verbs that supply causatives for these that is remarkably systematic, not what one might expect if, as Jamison says, one of the constructions was used to avoid ambiguity. Under her explanation, one should expect a more random distribution, the potentially ambiguous construction steadily losing ground. However, what is potentially ambiguous in isolation is not so in context, and speakers use sentences in contexts. For example, consider

(191) TS 6.6.2.2:
a. *yájña yajñáṃ gaccha yajñápatiṃ gacchéty āha* (TS 1.4.44.3) '(The Adhvaryu) says "Sacrifice, go to the sacrifice (i.e. Viṣṇu), go to the patron of the sacrifice".'
b. *yajñápatim eváinaṃ gamayati* 'He makes it go to the patron of the sacrifice.'

Taken out of context, (191b) could be ambiguous, but, as can be seen from the anaphoric pronoun *enam* (acc. sg.), this is embedded in a context, where it is unambiguous.

4.2 In an earlier study of Sanskrit causatives, Thieme also tried to demonstrate that early Vedic had no true causatives, formed to transitive verbs. I agree with Thieme (1929:19) that sentences such as (7b) result from influence of causatives formed to intransitive verbs, that is, from an

extension of the double accusative construction. Moreover, Thieme's treatment of the issues shows the mastery of Vedic and Pāṇinian materials so typical of this scholar. Nevertheless, I regret I cannot consider his arguments fully cogent. In essence, Thieme tries to show that all verbs which formed derivates in *-i* that one might call causative were intransitive in early Vedic. He takes up (1929:21-2) the causative construction for verbs of movement etc. in the Sanskrit of Pāṇini's time (see 1.1, [A]), refers to Vedic passages where this is found, and then says (1929:22) concerning the verbs in question: 'Zu allen diesen Verben ist ein Intransitiv denkbar, das ausdrückt, dass die vom Verb bezeichnete Handlung am Subjekt in Erscheinung tritt. Dieses Intransitiv kann entweder durch das blosse Medium ausgedrückt werden (*vardhati* "stärkt," *vardhate* "wird stark") oder--was besonders häufig ist--durch das Medium mit *-yá* (*mṛjyáte* "wird sauber"), oder durch das Verbaladjektiv mit *-tá* (*vṛddhá* "einer der gewachsen hat," gegenüber *pītá*, das auch heissen kann "einer der getrunken hat"). Tritt zu diesem Intransitive der Agens im Instrumental, so sprechen wir von *Passiv*. Z.B. *harṣate* "freut sich," aber X.30.5 *yābhiḥ ... hárṣate* "durch die er erfreut wird." *mṛjyate* "wird sauber, lauter," z. B. in IX.86.6 *yádi pavítre ádhi mṛjyáte háriḥ*, aber IX.76.4 *yáḥ sūryasyāsíreṇa mṛjyáte* "der geläutert wird durch . . ."' [Emphasis Thieme's].

The above actually involves two points. Let me begin with the claim that an intransitive middle is conceivable for all the verbs in question. The verb *gam* regularly takes middle endings in Pāṇini's time if it is accompanied by the preverb *sam* (Pāṇini 1.3.29: *sámo gamyṛcchibhyām* [*kartari* 14, *ātmanepadam* 12]): *saṃ gam* 'come together'. This is so also in earlier Vedic, as can be seen from examples such as (158) and RV 1.119.3a: *sáṃ yán mithāḥ paspṛdhānāso ágmata* '. . . when they came together competing one with the other'. And *úd gam* 'go up' also is found with ātmanepada affixes (e.g. RV 8.69.7d). However, such uses of the middle of *gam* can hardly account for causatives like *gamayāmasi* in (77), at least not in the way required by Thieme. It is true that *śrāvayati* 'make to be heard' is paired with a middle like *śṛṇve* 'is heard', but *śrāvayati* 'makes to hear' cannot be considered merely a transitive related to an intransitive middle. It must be paired with *śṛṇoti* 'hears' (see 2.2, 2.6). Similarly, though *cetayati* 'makes to be perceived, to appear' is a transtive-causative related to *cite* 'is perceived, appears', *cetayati* 'makes to perceive' must be coupled with *cetati* 'perceives'. It is also unacceptable to consider that *darśayati* 'makes to see' is merely a transitive to a middle *dṛśyate* 'is seen, appears', and for the same reason. Under Thieme's thesis, *darśayati* would have to be a transitive to *dṛśyate* meaning 'appears', but a transitive of the type *cetayati* with an accusative denoting what is made to appear and a dative signifying one to whom this is shown, as in (103), simply will not account for sentences like (92), (93). As shown in 2.12, *pā* 'drink' is transitive. This verb usually has active inflexion except in passive usage. It is true that early texts also have non-passive forms of this verb; for example, RV 10.114.7cd: *ā́ yāvānaṃ tīrthám kā́ ihá prá vocad yéna pathā́ pra píbante sutásya* 'Who will proclaim here the Āpnāna ford, the path on which (the gods) drink the pressed Soma?' Note, however, that *pra píbante* occurs with *sutásya* here, and is thus syntactically like *píba* in (95). It is also true that such middle forms of *pā* can be used absolutely, as in RV 10.135.1b: *devaíḥ saṃ píbate yamáḥ* '. . . Yama drinks with the gods'. However, the active of *pā* is similarly

used, as in RV 1.23.18b: *yátra gávaḥ píbanti naḥ* '. . . in which our cows drink'. Consequently, it is not possible to account for the causative *pāy-i* in the manner Thieme suggests. The verb *vac* too is transitive (2.12). To be sure, there are non-passive middle forms of this verb, but they do not conform to the requirements of Thieme's thesis, since they are regularly transitive. For example, in RV 4.5.11a: *ṛtáṃ voce námasā pṛcchyámānaḥ* 'I will tell the truth when asked with a show of obeisance' the middle form *voce* (1 sg. aor. inj.) occurs with the object accusative *ṛtám* 'truth'. No matter how conceivable it might at first appear, then, the fact is that saying derived bases such as *gam-i*, *śrāv-i*, *daré-i*, *pāy-i*, *vāc-i* originally were nothing more than transitive derivates corresponding to intransitive middle forms cannot be justified from the evidence.

Thieme makes his second point by bringing in sentences of type (56) *odanaḥ pacyate* 'The rice is cooking', when he speaks of *mṛjyate* used with or without an agentive instrumental. This sentence type does occur in Vedic. For example, TS 5.1.9.2-3: *sā yád bhidyétārtim árcched yájamāno hanyétāsya yajñáḥ* 'If it (the ukhā pot) were to break, the patron of the sacrifice would undergo ruination, his sacrifice would perish' has *bhidyéta*, *hanyéta*, optative forms corresponding to intransitive middles *bhidyáte* 'breaks', *hanyáte* 'perishes'. However, bringing in this type does not overcome the objections set forth above. In fact, this raises additional problems. The sentence type (56) regularly has verbs which can and do take object accusative complements and have true passives. The feature particular to this type is that an external agent performing an act on an object may not be specified, and the verb behaves like an intransitive verb. However, not all verbs enter into this construction, and verbs of movement such as *gam* do not. Moreover, if a form like *pacyate*, *bhidyate*, *hanyate* is used passively ('is cooked', 'is broken', 'is killed'), an external agent is signified by an instrumental. Consequently, under Thieme's thesis, one should expect only construction (70c) in causative sentences, not also (69c).

Of course, Thieme is well aware of the problem, and he anticipates Jamison's position, by arguing that verbs which have causatives with construction (69c) are actually intransitive. Thus, he says (1929:20) that *párāṃ ... paráāvátam* of (77) is an accusative of direction (Richtungsakkusativ), *vā́jam* of (132) an accusative of content (Inhaltsakkusative), so that the primitive bases *gam*, *ji* from which *gam-i*, *jāp-i* derive are intransitive. He does admit that accusatives used with verbs like *ṛkṣ*, *budh* are object accusatives. For he calls these transitive verbs (1929:20). Still, he hesitates to give them full transitive status, saying (1929:20), 'Transitive Verben der Art wie *ṛkṣ*, *budh* jedoch sind ihrem Ursprung nach intransitive Verben des Vorgangs, die mit einem freieren Akkusativ, dem Akkusativ *der Erstreckung*, konstruiert werden' [Emphasis Thieme's], and in a footnote (1929:20 n 2) he further suggests these verbs may originally have been construed with an accusative of goal (Akkusativ des Ziels), of direction. Later (1929:22), Thieme reiterates that *gam* takes an accusative of direction, a verb of perception an accusative of extent, though here he puts 'der Erstreckung' within inverted commas. Further, according to Thieme (1929:20-1), a genitive used with a verb of consuming like *pā* 'drink' is a partitive genitive, and an accusative construed with such a

verb is one of extent: 'Ähnlich liegt der Fall bei *pā* "trinken", das mit dem *acc.* aber auch mit dem *gen. part.* konstruiert werden kann, wenn nähmlich die Handlung sich nur auf einen Teil der bezeichneten Sache *erstreckt*' [Emphasis Thieme's]. The above does not withstand scrutiny. One must obviously recognize an accusative of goal, which alternates with other case forms (2.1), as one has to distinguish, with respect to verbs of movement, an accusative denoting a road taken and one denoting a road to which one goes, as in (159), (160). Moreover, one might claim that *ṛchati*, *ṛtāḥ* in examples such as (165), (167) are forms of a verb that has undergone semantic specialization, meaning 'afflict, attack', so that it would not be strictly a verb of movement. However, there is no escaping the fact that an accusative of goal construed with *gam* is treated in the same manner as an object accusative, witness (158), and whatever one says about *ṛchati*, the fact remains that it is transitive. Hence, *gam-i*, *arp-i* can be nothing other than causatives to transitive verbs. Whatever be one's opinion concerning the possible origins of accusatives used with verbs of perception, it is difficult to see how these could be equated with adverbial accusatives in examples such as (114), (185), which can rightly be termed accusatives of extent. Similarly, though a genitive like *sómasya* in (95) be labelled partitive, there is no obvious semantic difference between (94) and (95), and there is no justification for claiming the accusative *sómam* of (94) is of the same type as accusatives of extent found elsewhere. It is also difficult to justify the claim that *ji* and *jāp-i* in examples like (132)-(134) are construed with an accusative of content, a cognate accusative. A phrase such as *jítiṃ jigyúḥ* in ŚB 3.2.2.2: *yajñéna vai devā imā́ṃ jítiṃ jigyuḥ* 'The gods won this victory through the sacrifice' clearly contains a cognate accusative *jítim* 'victory'. This is not the case in RV 8.76.4ab: *ayáṃ ha yéna vā́ idáṃ svàr marútvatā jítam* 'This (is Indra,) by whom, with his troop of Maruts, this sun was won', since *svàr* 'sun' is totally unrelated to the base *ji* of the participle *jítám* 'won conquered', nor is *vā́jam* in (132)-(134) a cognate accusative. It might well be that phrases like *vā́jam jaya svàr jítam* could represent an extension of an original construction with a cognate accusative. However, this cannot justify arguing that *ji*, incontrovertibly transitive in Vedic, is somehow not a true transitive verb in examples like (132)-(134) because what is spoken of as won is a thing and not a person conquered.

4.3 Two major attempts to show that early Vedic did not have true causatives, formed to transitive verbs, thus fail.[24] Given what has been shown, then, I consider inescapable the conclusion that earliest Indo-Aryan did indeed have true causatives, derived from transitive verbs of particular semantic groups.

5.0 The early Iranian evidence is, as one might expect, more sparse than that of Indic. However, as one might also expect, it does not disagree with the Indic evidence.

As the Indo-Aryan verb *gam* occurs with an accusative, a locative, or a dative denoting a goal (2.1), so too does the Avestan verb *gam*. Thus, *paṛā hyaṯ ... jimaitī* 'before ... come' occurs with the accusative *mā* 'to me' in

(192) Y 48.2ab: *vaocā mōi yā tvəm vīdvå ahurā/ parā hyaṯ mā yā məng pərəθā jimaitī* 'Lord, tell me what you know, before the retribution in your mind come to me.'

and in Y 49.1c, *gaidī mōi* 'come to me', *gaidī* (2 sg. imper.) is construed with the dative *mōi* 'to me'. Further, in

(193) Y 51.15b: *garō dəmānē ahurō mazdå jasaṯ pouruyō* 'The Wise Lord came first into the house of chant.'

jasaṯ occurs with the locative *garō dəmānē* 'house of chant'. From this verb is derived the causative *jāmayeiti*, as in

(194) Yt. 17.20: *raēkō mē haca aŋhå zəmaṯ vaŋhō kərənaoiti yō mąm aēvō jāmayeiti yō spitāmō zaraθuštrō* 'He makes it better that I leave this earth, Spitama Zarathushtra, who alone makes me leave.'

where it is used with one accusative, *mąm* 'me'.

The verb *fra par* (preverb *fra*) 'cross over' occurs with an accusative denoting what one crosses; for example,

(195) Y 46.10e: *frō tāiš vīspāiš ainvatō fra frā pərətūm* 'With all these would I cross the bridge of the elector.'

where *fra frā* (1 sg. aor. subj.) is used with the accusative *pərətūm* 'bridge'. On the other hand,

(196) Y 19.6: *yasca mē aētahmi aŋhvō yaṯ astvainti spitama zaraθuštra baɤəm ahunahe vairyehe ... frā vā ... srāvayāṯ ... θrišciṯ tarō pərərtūmciṯ hē urvānəm vahištəm ahūm frapārayeini* 'Spitama Zarathushtra, I will make his soul thrice to go across the bridge (of the elector), to the best life, who in this corporeal world recites out loud the Ahuna Vairya prayer.'

contains the causative *frapārayeini* 'I will make to cross' in construction with *tarō pərətūmcit* 'across the bridge' and two accusatives: *urvānəm* 'soul' and *vahištəm ahūm* 'the best life (paradise)', of which the latter denotes a goal. In addition, the causative *nipārayeinti* 'they cause to come, bring' in

(197) Yt. 17.54: *mā ciš aŋhąm zaoθranąm vindita yā māvōya nipārayeinti* 'No one shall have part in these oblations which they bring me . . .'

is construed with the accusative *yā* '(the oblations) which' and the dative *māvōya* 'to me'.

An object of perception can be denoted by an accusative or a genitive in Indo-Aryan (2.2), and this is so also in Iranian. Thus, the Avestan cognate of Sanskrit *śru*, the verb *sru* (2 sg. opt. *surunuyå*) takes the accusative complement *yasnəm* 'prayer' in Y 68.9: *surunuyå nō yasnəm ahurāne* 'Ahurānī, hear our prayer', but in Yt 10.32: *surunuyå no miθra yasnahe* 'Mithra, hear our prayer', the same verb occurs with the genitive

yasnahe. In addition, middle forms of *sru* are used in the same manner as those of *śru*, meaning 'be heard', 'be heard of' (2.6). For example,

> (198) Y 34.12ab: *kaṯ tōi rāzarə kaṯ vaŝt kaṯ vā stūtō daṯ vā yasnahyā/ srūidyāi mazdā frā vaocā* ... 'What is your command, what do you wish in praise or prayer of veneration? Wise one, proclaim it that one might hear . . .'

contains the medio-passive infinitive *srūidyāi* 'that . . . be heard' (cf. Benveniste 1935:76), and Y 46.14b: *kə vā frasrūidyāi vaŝtī* 'Who wants to be famed?' has *frasrūidyāi* 'to be heard of, famous'. Further, the finite form *sruyē* (1 sg. mid.) is used in this sense in

> (199) Y 33.7ab: *ā mā idūm vahištā .../ aŝā vohū manaŋhā yā sruyē parə magaonō* 'Come to be, best ones, with Truth and Good Conception, that I be heard before the priests.'

Corresponding to this medio-passive, there is a causative *srāv-i*; for example, Y 49.6cd: ... *yaθā ṯ srāvayaēmā taṃ daēnaṃ yā xšmāvatō ahurā* '. . . how we may make heard this your religious insight, Lord'. This causative is also used specifically with reference to reciting sacred texts, as in (196).

The Avestan verb *vard* is cognate with Sanskrit *vṛdh* (2.7) and used in the same meanings. First, it means 'grow, increase':

> (200) Yt. 13.68: *xᵛaēpaiθe nō daiŋhuš fraδataeca varəδātae ca* 'Our land is to flourish and grow.'

In addition, the active (3 sg. *varədaitī*) is transitive, meaning 'make grow'; for example,

> (201) Y 28.3: *yə̄ vå aŝā ufyāni manascā vohū .../ mazdąmcā ahurəm yaēibyō xšaθrəmcā .../ varədaitī ārmaitiš ā mōi rāfəδrāi zavāŋg jasatā* 'Come to my calls, to help me, I who will praise you in chant with truth: Good Conception, the Wise Lord, and ruling power, for whom Devotion makes grow their ruling power.'

Moreover, the derivate *vard-i* is used in the same sense as the active:

> (202) V 2.4: *āaṯ mē gaēθå varəδaya* 'Then, (Yima,) make my world increase.'

Similarly, the primitive verb *paiti.irinaxti* 'abandons' and the derived verb *paiti ... raēcayeinti* 'abandon' are equivalent: Yt. 14.47: *kə̄ rašnūm paiti.irinaxti* 'Who abandons Rašnu?' Yt. 10.41: *paiti θrātāra yazata ... rasmanō raēcayeinti* 'The gods who protect them abandon the regiments'. This too is paralleled in Indo-Aryan (2.8, 2.10).

6.0 Since the early Iranian data are in harmony with those of Indic, I think it safe to say that the causative system found in earliest Vedic is inherited from Indo-Iranian. Indeed, even if there were no Iranian evidence, one should be led to this conclusion. For, as has been shown,

the trend within Indic is for causative formations to be extended from an earlier limited domain, so that the early Vedic system, in which true causatives were formed to transitive verbs of particular semantic groups, could hardly be an Indo-Aryan innovation. What I call the old system, then, had the following network of contrasts.

I$_1$. Verbs of movement, perception, consuming, and verbs meaning 'say' or similar: For these, there was a three-way contrast among transitive, passive, and causative (2.1-2.4, 2.12). In addition, a causative could be related to either an active or a passive, so that two causative constructions could contrast, meaning '*C* causes *A* to *V B*' and '*C* causes *B* to be *V*-ed by *A*' (2.6).

II$_1$. Other transitive verbs: For these, there was a two-way contrast between transitive and passive. The contrast between causative and transitive did not obtain (2.8, 2.10).

III. Intransitive verbs: There could be no contrast between transitive and passive, nor was there one between transitive and causative (2.5, 2.7, 2.10).

In the innovating system, there is a contrast between transitive and causative for all transitive verbs. Moreover, by the time of Pāṇini the contrast between causative sentences meaning '*C* causes *A* to *V B*' and '*C* causes *B* to be *V*-ed by *A*' was being obliterated. However, the old system leaves its traces in syntactic dintinctions:

I$_2$. Verbs of movement, perception, etc. (*I*$_1$ above) as well as intransitive verbs: Causatives of these enter into construction (69c), in which an accusative denotes the agent of the non-causal act.

II$_2$. Other transitive verbs: Causatives of these regularly enter into construction (70c), in which an instrumental refers to the agent of the non-causal act.

The old system leaves another trace in that some derived verbs such as *mārj-i*, *vār-i* are also used without causative value (1.4, 2.10).

My theses, then, are these. First, earliest Indo-Aryan had true causatives, formed to verbs of particular semantic groups (*I*$_1$ above), and this state represents an inheritance from Indo-Iranian. Secondly, there took place in Indo-Aryan at a time certainly prior to Pāṇini a series of syntactic and morphologic changes, involving middle and passive forms as well as causatives, which led to the generalization of causative formations (3.1-3.3).[25]

ABBREVIATIONS

Ait.Ār.: Aitareyāraṇyaka; Ait.B: Aitareyabrāhmaṇa; AV: Atharvaveda; KBU: Kauṣītakibrāhmaṇopaniṣad; KS: Kāṭhakasaṃhitā; KŚS: Kātyāyanaśrau-

tasūtra; MS: Maitrāyaṇīsaṃhitā; RV: Ṛgveda; ŚĀr.: Śāṅkhāyanāraṇyaka; ŚB: Śatapathabrāhmaṇa; TB: Taittirīyabrāhmaṇa; TS: Taittirīyasaṃhitā; V: Vidēvdāt (Vendidād); VS: Vājasaneyisaṃhitā; Y: Yasna; Yt.: Yašt.

NOTES

*This is a revised version of the Herman Collitz Lecture delivered at the 1978 Linguistic Institute on July 25, 1978. An abbreviated version is also to appear in the Golden Jubilee Volume of the Vaidika Saṃśodhana Maṇḍala, Poona. I with to thank H. H. Hock and Bh. Krishnamurti for helpful discussions on the matters treated here and G. Dunkel for valuable comments on an earlier draft. I am also grateful to the National Endowment for the Humanities for support I received while writing this.

[1]For a brief presentation of Pāṇini's derivational system, see Cardona 1976b (pp. 146-8 on causatives in particular).

[2]I cite such isolated forms in their pre-pause variants, with -*ḥ* etc.; different ones may appear in sentences cited, due to phonologic environments.

[3]I say 'of itself' because combinations such as *adhi ās* 'remain in, rule over' are transitive.

[4]Thus Pāṇini, who treats the rice spoken of in these sentences as agent (Cardona 1974:241-3). Forms of the type *pacyate* 'is cooking' are not used for certain verbs; e.g. *dugdhe* 'gives milk' (*duhyate* 'is milked'), *namati* 'bends'.

[5]In earlier Sanskrit, a locative is also used; see (76). In the Sanskrit of Pāṇini's time, a verb such as *viś* 'enter' could still take a locative complement, but not if used with the preverbs *abhi*, *ni*; *abhi viś*, *ni viś* govern accusatives.

[6]The use of a genitive in construction with certain verbs is more widespread in earlier Sanskrit (see 2.2, 2.3). I have omitted many details concerning locative and genitive complements with particular verbs in later Sanskrit.

[7]Thus Pāṇini; see Cardona 1974:289 n 49.

[8]That is, the nominal N_2 can be followed by different endings, as shown.

[9]N_2 in b is followed by a nominative ending if it has an accusative ending in a; otherwise, the same ending occurs with N_2 in both.

[10]V-*i* is a causative base derived from the primitive verb V of a, b.

[11]The last section of the dhātupāṭha includes a rule (*ā dhṛṣād vā*)

according to which ṆiC is optionally added to bases which follow, up to and including *dhṛṣ* (see Pathak-Chitrao 1935:668 [immediately preceding verb 1807]), without causative value; *arc-i* etc. given above are among these verbs.

[12]See Cardona 1974:241. Forms of the type (50) *āsi*, (58) *apāci* have the affix CiṆ, which in Pāṇini's derivational system is introduced before the ending *ta*. This ending is then zeroed. Passive forms other than the ones noted simply have middle endings.

[13]But in the sense 'protect', *bhuj* has agential forms with parasmaipada affixes, *bhunakti* etc.

[14]See Cardona 1976b:148 n 41.

[15]I have of course given only a bare outline of the distribution, and cannot consider more details here.

[16]Here again, I must omit many details. Note nevertheless that *bodh-i* ([2b] *bodhayati*) does not have a middle under the condition given. Nor does a causative like *bhoj-i* ([3b] *bhojayati*), to a verb of consuming, though *pāy-i* ([96] *pāyáya*) can have a middle.

[17]I do not imply, by speaking of reflexives, that one should consider a reflexive pronoun implicit in such sentences; see Cardona 1976b:148, 156-7.

[18]For Iranian parallels, see 5.

[19]In Vedic, genitives construed with *śru* regularly denote persons to whom one listens, but in Iranian a genitive can also signify a thing listened to; see 5.

[20]That a comparable passage (RV 7.60.6b) has *acetásam* (acc. sg.) 'lacking insight' does not affect the grammatical relation noted.

[21]Cf. also *āvir bhū* 'become apparent', *āviṣ kṛ* 'make apparent', construed with a dative denoting one to whom something becomes or is made apparent.

[22]This is true also of modern Indo-Aryan, where distinct causative constructions still exist for different types of transitive verbs. On the other hand, in major languages such as Gujarati and Hindi the double accusative construction is not used. Instead, an indirect object form denotes the agent of the non-causal act. I cannot discuss this or facts from middle Indo-Aryan in the present paper.

[23]I have not discussed here the aorist type which is regularly paired with causatives, the reduplicated aorist, but this does not affect my argument.

[24]I have concentrated on Jamison's and Thieme's views because these are the only scholars who have truly attempted to justify their positions.

Earlier, Gaedicke (1880:277) proposed an explanation which was refuted by Thieme (1929:20). Delbrück (1888:226) suggested that an instrumental appears in a casuative construction because the subject appeared so much to be the main participant that the second agent was demoted to being a means. This fails to explain why double accusative constructions also were used. Most recently, Haudry (1977:382-9) has treated causative constructions in a manner which is vague, ill-founded, and fails to account for the chronology of different constructions. I discuss his view in a review of his book, to appear in Kratylos.

[25]In order to avoid ending on a note of speculation, I have not gone on here to discuss the question of causatives in a broader Indo-European context, though I plan to take this up. For the moment, let me only note that there is sufficient evidence from subgroups such as Germanic and even Italic to suggest that in Proto-Indo-European too causatives may have been semantically restricted, at least so far as concerns morphologically derived causatives. In addition, pairs such as *mārṣṭi*, *mərjayati* (2.8, 2.10, 5) doubtless represent a merger of forms that originally belonged to distinct categories, which brings up the question of iteratives. For the instant, I think it sufficient to emphasize that Indo-Iranian definitely had true causatives formed to transitive verbs of particular groups, so that any future study of the Proto-Indo-European type **gʷom-ey-e-t-i* (Skt. *gəmayati*, Av. *jāmayeiti*), **louk-ey-e-t-i* (Skt. *rocayati*, Av. *raocayeiti*, Lat. *lucere*, as in Plautus, Casina 118 *lucebis ... facem* 'you'll light the bridal torch') must start with this as a datum; cf. also Lat. *disco* 'learn', *doceo* 'teach', a causative found in a double accusative construction, and a causative such as Gothic *dragkjan*, which corresponds to the transitive *drigkan* 'drink'. Moreover, one must keep in mind that Proto-Indo-European could well have had periphrastic causatives such as shown in 2.11.

REFERENCES

Abhyankar, K. V. 1962. The Vyākaraṇa-Mahābhāṣya of Patañjali, ed. by F. Kielhorn. 3rd ed., vol. 1. Poona: Bhandarkar Oriental Research Institute.

Benveniste, Emile. 1935. Les infinitifs avestiques. Paris: Adrien-Maisonneuve.

Cardona, George. 1974. Pāṇini's kārakas: agency, animation and identity. Journal of Indian Philosophy 2. 231-306.

------. 1976a. Pāṇini, a survey of research. The Hague: Mouton.

------. 1976b. Some features of Pāṇinian derivations. History of Linguistic Thought and Contemporary Linguistics ed. by H. Parret, 137-58. Berlin-New York: de Gruyter.

Delbrück, Bertold. 1888. Altindische Syntax. Repr. 1968. Darmstadt: Wissenschaftliche Buchgesellschaft.

Gaedicke, Carl. 1880. Der Accusativ im Veda. Breslau: Koebner.

Haudry, Jean. 1977. L'emploi des cas en védique. Introduction à l'étude des cas en indo-européen. Lyon: Editions L'Hermès.

Jamison, Stephanie. 1976. Functional ambiguity and syntactic change:

the Sanskrit accusative. Papers from the Parasession on Diachronic Syntax, 126-35. Chicago: Chicago Linguistic Society.
Kuryłowicz, Jerzy. 1956. L'apophonie en indo-européen. Wrocław: Polska Akademia Nauk.
------. 1964. The Inflectional Categories of Indo-European. Heidelberg: Winter.
Pathak, Shridharshastri and Siddheshvarashastri Chitrao. 1935. Word Index to Pāṇini-sūtra-pāṭha and Pariśiṣṭas. Poona: Bhandarkar Oriental Research Institute.
Speyer, J. S. 1896. Vedische und Sanskrit-Syntax. Strassburg: Trübner.
Thieme, Paul. 1929. Das Plusquamperfektum im Veda. Göttingen: Vandenhoeck und Ruprecht.

CONSERVATIVE WORDS AND RADICAL SENTENCES IN THE SEMANTICS OF INTERNATIONAL *POLITICS*

Charles E. Osgood
University of Illinois

Albert Einstein, with the foresight characteristic of true genius, had this to say at the dawn of the nuclear age: 'The unleashed power of the atom has changed everything except our ways of thinking. Thus we are drifting toward a catastrophe beyond comparison. We shall require a substantially new manner of thinking if mankind is to survive.' It will be my thesis that a new manner of thinking about our world will first require a revolution in how we talk about it.

I shall try to demonstrate that certain principles of psycholinguistics have relevance to the science, and art, of politics--particularly international politics. Psycholinguistics is that facet of the human sciences which deals with relations between the characteristics of messages and the states of the (human) organisms which exchange them. We shall be concerned primarily with the semantic aspects of these relations--with the meanings of word forms, with the rules by which words are combined into phrases and sentences, with the effects of one's utterings and scribblings upon others, with the semantic and valuative constraints upon what we ordinarily say with our words and sentences, and with how we can get out of this bind by some calculated rule-breaking.

You may well be wondering already just what all this could possibly have to do with politics, particularly international politics. To satisfy this wonderment is, or course, the task I have set for myself. Throughout this talk I shall introduce each new notion with illustrations drawn from ordinary language and only then move into illustrations and analyses in the realm of international politics. I adopt this strategy for two reasons: first, because we are all most familiar with the ordinary language we use in communicating with family, friends, and storekeepers, and the examples are therefore more compelling; second, because examples drawn from contemporary international relations, necessarily reflecting my personal political values, are liable to be controversial and hence lose their force.

Personification of international relations is a common practice of social scientists and statesmen, just as it is among laymen. It is even to be found regularly in the writings of those scholars who would be the first to deny, as a matter of disciplinary principle, that laws of individual behavior apply to social groups. Such implicit personification --and consequent uncritical use of folk psychology in the interpretation of international relations--is both deceptive and dangerous. But there is a paradox here: although one may deplore such naivety, if personification is in fact common practice in thinking and decision-making in

international affairs, then it must be considered part of the subject matter of political science and made as explicit as possible. I shall assume that principles of semantics discoverable at the individual level can be transferred to the national level--because individual humans do participate in the 'behaviors' of nations, because certain individuals do play key roles in the 'decisions' of nations, and because principles of information-processing (perceiving, interpreting, decision-making, executing) do have their analogues at all levels.

CONSERVATIVE WORDS

Words are conservative. Although they adapt to changes in the real world, they do so very slowly if left to themselves. Assuming a lawful and immutable relation between words and things, we humans keep trying to force things to conform to the meanings their words have acquired in the past. Thus were the semantics of World War I applied in preparing for World War II, and the French built a completely useless Maginot Line and a false sense of security behind it. So do we seem to be preparing for World War III, despite desperately not wanting it, by applying the semantics of World War II.

Wrapped up in the whole process as we are, it is difficult to perceive the lag between our language and the world it is supposed to describe and interpret. Occasionally a dramatic failure of events to conform with our expectations may jolt us into awareness that our world is changing--a rash of race riots or a score of people suffocating in the smog of an American city. The brighter among us may note that there is a constant conservative bias in our predictions about the future--sputniks appeared in our skies and the Chinese tested deliverable nuclear weapons long before these things were supposed to happen.

In periods of history when the rate of cultural change was slow, words could almost keep up with things and their mismatch was barely perceptible. But in the present age, with its exploding population and its exploding technology--which means increasing human interaction and accelerating cultural change--our semantic maps of the real world become outmoded more and more quickly. Add to this the fact that most decisions in today's world are being made by men in their fifties, sixties, seventies, or even eighties--men whose maps were outlined at least 30 years ago--and the language gap threatens to become an abyss. Led by old men reading old maps, we are trying to find our way through the Wonderland of the Twentieth Century, having more contact with the phantasy of words than with the reality of things.

The gap getween word and thing also increases with the remoteness of things from immediate, individual experience. As one moves outward--away from the intimacy of family, possessions, and the ordinary round of living, toward the community, the nation, and the world--the map becomes a less and less reliable guide to the territory. For one thing, we become less able to correct our verbal maps against the facts of the terrain. For another, the meanings of our words become increasingly dependent upon the words of others. The words with which we talk about the most critical issues of our time--nuclear overkill, arms control,

population control, and Palestinian Commandos--are what in the Business of Semantics are termed 'assigns'. Their meanings are literally assigned by association with other words put together by other people (usually in the mass media) rather than from direct experience with things and events. We have immediate knowledge about what a knife can do: our knowledge about what a nuclear missile can do is rather remote and gutless. Thus, irrationally, knowing that one's city is targeted by a 20-megaton missile some ten thousand miles away is not as threatening psychologically as knowing that a madman with a carving knife is loose in the city streets.

The Power of the Word

Words may be conservative, but they can also be powerful. Wherein lies the power of the word? It obviously doesn't lie in the noises and squiggles per se. It lies, rather, in a very remarkable relation between these physical manifestations, called 'signs', and certain processes in language-users, called 'meanings'. This is the representing relation. Words come to evoke in their users some distinctive representation of the things referred to, and these representations can be manipulated symbolically much more easily than the things themselves. Thus we can behave--appropriately or otherwise--with respect to the not-here and the not-now. A Jimmy Carter can devise strategies in which nation-names move about his mind like chessmen on a board; words about things remote can give him hope (President Nasser makes a peace initiative by going personally to Israel to talk peace in the Middle East) or they can give him fear (the possibility that Communist Cuba and Russia may threaten the dominance of the West in Africa and threaten 'balance of power' stability).

Exactly what noises and squiggles will be used to represent what things and events is largely arbitrary--which is one reason why human languages are mutually unintelligible. When one is born into a given language-culture community, he unwittingly assumes a social contract to use noises and squiggles the way others in his community do--to use the noise *man* to refer to adult male humans, for example. But once the myriad contracts of lexical and syntactic usage have been agreed upon, the relation of meanings to things is not arbitrary--which is one reason why human languages are mutually translatable.

The meaning of a word can be conceived as a simultaneous bundle of distinctive semantic features. These features represent those differences in reactions to things and events which, in human experience, have been found to make a difference in adjustment to the physical and social environments. Thus the representing or symbolizing process utilizes a kind of code for highlighting what properties of things and events have been found to be important in experience. Recent studies in comparative psycholinguistics are beginning to make it clear that human groups, regardless of differences in language or culture, tend to use very similar systems of semantic features--as if, in general, they have found the same kinds of distinctions important to make.

In some of our own research, for example, we have been finding that humans highlight certain properties of interpersonal behavior, but not others, as semantic features in their interpersonal verbs. What is also

interesting is that they use many of the same verbs with the same features to talk about international relations, as if nations were people. There is an Associative-Dissociative feature (*help* vs. *hinder*, *protect* vs. *attack*), an Ego- vs. Alter-oriented feature (*exploit* vs. *aid*, *manipulate* vs. *advise*), a Superordinate-Subordinate feature (*dominate* vs. *submit to*, *lead* vs. *follow*), a Future-Past feature (*promise* vs. *apologize*, *threaten* vs. *retaliate*), and so on. Needless to say, there is considerable danger of misrepresentation when we apply, willy-nilly, the semantics of interpersonal relations to the things and events of international relations.

Since the semantic features on which words are coded can never exhaust the properties of things which might be represented, it is inevitable that words must be abstractions. In this sense, words are caricatures of things. Although this is most obvious for big, impressive symbols like *Colonialism* or *Black Power*, it holds also for little ordinary words like *dog* and *pencil*. It is precisely in this selective abstraction from reality that the power of words lies. Those properties of things that are semantically coded are sharpened in thinking; those properties of things that are not so coded are blurred in thinking. A man can freeze to death for lack of kindling wood and yet be found with a pocketfull of wooden pencils--because pencils are coded 'to write with', not 'to make fires with'.

What facets of reality are sharpened or leveled by words depends on what properties of things have made a difference in the past use of the language. This is the conservatism of words, and it can make us attend to features of events which are now irrelevant and be oblivious to features which are now critical. Bound by an antiquated semantics of Power Politics, diplomats keep wary eyes on the shifting patterns of alliances and scurry about trying to maintain *mutual security pacts*--NATO's, SEATO's, CENTRO's, and PREPOSTRO's. The semantic coding of *mutual security pact* simultaneously fixes thought on the solidarity of one group and on its antagonism toward another. But the differences within the groups so segregated and the similarities between their particular members are obscured. Within the so-called Free World there are members which, by most criteria, are more similar to members of the so-called Communist Bloc, and vice versa.

The semantic features of words may either define or imply. For English speakers, the term *father* includes Male, Adult, and Parental among its defining features, and Dominance and Goodness among its implicative features. Any human we call a *father* must be an adult male who has progeny, or else we are breaking our linguistic contract, but obviously he need be neither good nor dominant. Substitution of implicative for defining features--that is, the use of emotive language--is the trademark of propaganda. Buck teeth and a spider body obviously did not define a Japanese human, but it did make him more killable in World War II. Just as the same odor may be called an *aroma* or a *stench*, depending on how one wants the listener to feel about it, so may the same guerrillas be called *freedom fighters*, *rebels*, or *terrorists*, depending upon how one wishes the listener to feel about them.

Words of Power

Among the few remaining places where myth and fantasy still have power for modern man are advertising and international relations. Although we cannot eat words, smoke words, or shave with words, the advertising industry makes a valiant effort to achieve that end; incantations, spells, contagious and analogic magic are part of the trade. International politics has its incantations ('the right to self-determination'), its spells ('we have a commitment'), its contagious magic ('a threat to freedom in Asia is a threat to freedom in Podunk, U.S.A.'), and its analogic magic ('another Munich'). Both advertising and international relations have everything it takes for myth-making--remoteness from the individual, conflict between ultimate Goods and Evils, and near-omnipotence of the storytellers. And, in both, the myths created do serve many social functions: they provide an illusion of understanding, simple prescriptions for ordering a complex world, and the promise that, if the illusions are believed and the prescriptions followed, everyone will live happily ever after. In the political arena, at least, the storytellers are often their own most gullible audience.

The words of international power politics are typically analogic. There are analogies from interpersonal relations; nation-states become *We*'s and *They*'s who assume military *postures*, who glare at each other *eye-ball to eye-ball*, who *harass* or *intimidate* each other, who *trust* or are *suspicious* of each other. There are analogies from 19-century physics, with *balances* of power or power *vacuums*, with *slippery slopes* or *rows of dominos* along which vague forces operate with some natural inevitability, with *attractions* and *repulsions*, and centers toward which power *gravitates*. Analogies have the same sharpening and blurring effects upon thinking that ordinary words do, only more so; they raise the feature analogized to a dazzling prominence which obscures other, often more relevant features.

The pseudonyms which pepper dialogue in the nuclear age carry this attenuation of word from thing even further, adding semantic features which do not exist in the thing and neatly countermanding features which do exist. The title CAMELOT conferred a romantic, even chivalrous, tang to an ill-fated U.S. Army project designed to study the causes of revolutions. To name an anti-ballistic missile system SAFEGUARD certainly must make its possessors feel more secure. A touch of nobility is added to raw power when intercontinental ballistic missiles are named THOR, JUPITER, ATLAS, ZEUS, and POLARIS--although I miss the ultimate in semantic deception which would be a missile named VENUS!

Acronyms--short-cuts which compress the forms of phrases into their initial letters--carry the attenuation of word from thing even further. They also put a great burden on human memory. Debates on the status of NATO and on the merits of SALT negotiations can be carried on with only remote reference to the things involved. They become 'things' of sorts on their own, things to be defended or attacked by partisans or critics with more reference to personal and political loyalties than to the complex events and arrangements in the real world they actually represent. When an acronym is devised before a thing is given an ordinary name, contact

with reality is liable to disappear entirely. A playful example from the argot of military space scientists is EGADS, which names the system used to destroy a malfunctioning missile after it has been launched--translation: 'electronic ground automatic destruct sequencer'!

Much of military jargon is designed to dehumanize things and thus protect the speaker or listener from feeling moral qualms. In *The Domesday Dictionary* of Kaplan and Schwerner (1963), for example, we find that the tertiary effects of nuclear blast are 'damage received by displacement of the biological target, or blast flight' and this is further analyzed into 'differential displacement' (which translates as violent separation of hand, arm, leg, head, etc.) and 'total displacement' (which really translates as being tossed into the next block in one chunk) . . . The military do have a way of juggling semantic features. We never bombed a simple village hut in South Vietnam, it was always a 'VC structure'.

Here are a few words of power to conjure with: The word *power* itself is often confused, by those who think they have it, with the word *force*. The meaning of *power*, unlike *force*, includes the feature of being able to produce a desired result. According to political scientist Murray Edelman, 'force signals weakness in politics, as rape does in sex.' If the use or threat of force serves to create the very outcome it was designed to prevent--for example, setting up military bases in Thailand resulting in a sharp increase in communist counteractivity--then it is selfdefeating and certainly not a demonstration of power.

Peace and *freedom* are nice words, but they are otherwise rather empty semantically. And being empty, they lend themselves to bizarre uses. Is *peace* a world spinning toward eternity triggered for mutual annihilation and kept from it only by mutual fear? This is what the phrase 'peace through military strength' really means today. The name of *freedom*, along with the resounding phrase 'law and order', is often used to justify its exact opposite--to rationalize subservience to authority and to condone suppression of the freedom of others.

The term *coexistence* does reflect an awareness of the irrationality of mutual nuclear annihilation--but what more than this does it mean to people? At an International Convocation on the Requirements of Peace, held in 1965 by the Center for the Study of Democratic Institutions, Paul-Henri Spaak of Belgium said: 'coexistence means only that you and we renounce war. For the rest we fight, trying to make the best demonstration we can that one system or the other is best.'

In is interesting that most participants in the conference used the phrase *peaceful coexistence* to describe this state of affairs. Surely, *competitive coexistence* would more aptly represent what they had in mind. There are many ways one can enliven the word *coexistence*, and by so doing point forcefully to quite different real-world possibilities: one could talk about *fearful coexistence*, pointing to the inherent nature of mutual nuclear deterrence; one could talk about *tolerant coexistence*, pointing to acceptance of differences without trying to change them, and even *co-operative coexistence*, pointing to co-operation rather than competition despite differences in systems.

In Defense of Ambiguity

Ambiguity can hurt. It does when the same word is used in the same context, but with different meanings for different people. Such is the case with *socialism* and *capitalism*--to say nothing about *democracy* as used by Americans vs. Russians, for example. But for most ordinary words, potential ambiguity is effectively eliminated by context. When someone says *duck* in the barnyard, I will not crouch as I will when the same word is uttered in a sand-lot ball park.

Proponents of One World Language as the salvation of Mankind often include elimination of ambiguity as one of the planks in their platform --one word one meaning! Not only would this increase the size of our working vocabulary inordinately--the little word *play*, for example, has some 40 different meanings, at least 15 of which are quite common--but a language free of ambiguity would be like sand in the fluid coupling of man-to-man and man-to-government. Some degree of ambiguity is necessary for a smoothly functioning society.

Political scientist Murray Edelman also makes the insightful observation that administration of laws governing ordinary social behavior tends toward a compromise between the letter of the law and the vagaries of human nature--too loose administration sets man against man and too strict sets man against government. When law is enforced as if it were a command rather than a virtuous generalization, around which a game can be played, it becomes a trial of force rather than a trial of wits.

In international politics there appears to be little tolerance of ambiguity. Negotiators strive for months to overspecify the inherently unspecifiable, to wring all possible ambiguities out of every statement in every language. At the Peace of Westphalia in 1648 it took six months for delegates to decide in what order they should enter and be seated in the negotiating chamber! How many weeks did it take to decide on the shapes and markings of tables for the U.S./Vietnam negotiations in Paris? International politics is more a trial of force than a trial of wits, and ambiguity is therefore more threatening than challenging.

A prime example of intolerance of ambiguity is reification of The Nation as the unit in international affairs. Since the flowing forests and oceans do not recognize these creations of the human mind, we put up boundary markers, erect walls and fortification, establish bordercrossing restrictions, try to impose language homogeneity within boundaries, define invisible territorial extensions into the seas and brightly color our maps in different hues so that children can learn just what is really where--all to reaffirm that The Nation is indeed a unitary thing like other things that have names. It then becomes easier to personify nations as Actors in a great global game, and harder to appreciate either the similarities across boundaries or the differences within.

Part of any semantic revolution must be an increase in the ambiguity of nation names, by using language which deliberately levels distinctions between nations and sharpens differences within them. This implies gradual dissolution of the nation-state as the prime political unit--which, surely,

is heresy. But in a nuclear age there may be more security in disunity; out of creative chaos there could come both a greater tolerance of ambiguity and more accurate semantic mapping of the complexities of the real world--paradoxical as that may seem. Fortunately, human languages contain the seeds for their own revolutions, so let us turn now from words to sentences.

RADICAL SENTENCES

We refer to our species as *homo sapiens*, but *homo loguens* would be more accurate. Other animals have been shown to symbolize the not-here and not-now to some degree--witness the way honey bees can represent the distance and direction of new food sources by their dance; other animals certainly manipulate the environment toward their own ends--witness the dams built by beavers. But no other animal really talks, in the sense of using a shared set of abstract rules for making propositions about the world. This is at once our great advantage as a species and, potentially, our downfall. Ability to propositionalize about things leads to control over them, but control over things without understanding and control over ourselves invites disaster.

Humans share many needs with lower organisms--needs to procreate, to feed, to avoid pain, for example. But the fact of propositional language creates uniquely human needs, particularly the need to know and the need to do something about what is known. When one can conjure with the not-here and not-now, one often assumes some responsibility toward it. Myths, religions and ordinary daily news are testimony to the need to know, or at least have some illusion of knowing, and knowing impels doing things in anticipation of events--making an offering at the shrine of one's God, planting crops at a time deemed propitious by omen or science, or even building a fall-out shelter in anticipation of a nuclear attack--which has now become sheer nonsense.

Valid propositions about the environment, either informal ('common-sense') or formal ('science'), have given humans considerable freedom from its harsh constraints. But scientific control of the environment has already reached the point where we now have a new kind of freedom--a freedom to, which is positively embarrassing in its riches: freedom to make deserts into gardens, freedom to manipulate our own numbers and even our own genetic makeup, freedom to bring any part of the world to any doorstep by satellite communication. The possibilities are staggering to the imagination, but choice among them demands fresh and apposite propositions about the world.

Rules for Creating and Understanding Sentences

We make propositions by combining words into sentences of a particular type. Propositional sentences make assertions about their topics, and they can be signaled by *I state that*, *I assume that*, *I claim that*, and the like --but we are rarely given this warning. Unlike words, propositional sentences can be tested for their truth value. Take, for example, the sentence *Tom is a thief*. One can ask significantly, *is it true that Tom is a thief?* One cannot ask significantly, *is it true that Tom?* or ask *is it true that thief?*

When propositional sentences come from prestigious persons and make assertions whose truth values cannot be readily tested--either because of the remoteness of the things referred to or because of the inherent ambituity of the concepts involved--their power of conviction is compelling. Faced with the assertion of a Pentagon expert that *a missile gap exists and threatens our security*, what is hapless Mr. Congressman to do but vote for a bigger defense budget--even though this assertion eventually proves to be false? Interestingly enough, resistance to such assertions, when it comes, is usually not so much based upon the truth-value of the assertions per se as upon the veracity of the speaker as a source. The peculiar power of sentences comes from the fact that they can assert something about the topic to be true--even this sentence, *what I say* (topic) *is not true* (commentary).

The meanings of words are constrained by the sentences in which they appear. Let us represent, for convenience, the meaning of a word as a string of pluses, zeros, and minuses--a strip-code for its bundle of semantic features--and then ask ourselves what may transpire when two or more words are forced to interact within phrases or sentences. (1) If the coding of one word is the same on a certain feature as another word with which it is pressed into syntactic combination, it will intensify that feature of meaning--as the adjective *sudden* does for the noun *surprise* in the phrase *sudden surprise*. (2) If one word adds distinctive codings where the other has zeros, then it serves to modify the meaning of the whole--as in the phrase *sudden anger* (one can also experience *slow, burning anger*). (3) If one word has codings opposed to those in another word on the same semantic features, then what are called 'semantic anomalies' are produced--as in *sudden complacency* or *meek contempt*. This 'feeling of incongruence' is a most significant human capability, of which I shall have more to say momentarily.

The assertion *Tom is a thief* momentarily punches some of the semantic codings of *thief* onto the topic of the sentence, *Tom*, thereby modifying the meaning of *Tom*--quite congruently for the arresting officer who caught him with his hand in the till but completely incongruently for Tom's Mother! Yet they both understand what the speaker intends, whether believing or not believing him. This momentary fusion of semantic features is essential for understanding phrases and sentences, but it would be disastrous if the change in word meanings were permanent. Fortunately, words tend to snap back into their normal semantic shape after being bent to the purposes of sentences. Otherwise our semantics would become a murky, meaningless shambles.

The influence of word upon word within sentences can be quite subtle. In English the word *neighbor* is not coded for sex. Yet if I say *My new neighbor is pretty*, you automatically assume that I am referring to a female, and if I say *My new neighbor is handsome*, you assume that it is a male. The sex codings of *pretty* and *handsome* are being momentarily conferred upon neutral *neighbor*. Sentences graced by the phrase *most people* carry a sense of ubiquity that far exceeds its casual method of calculation. Use of the phrase *Administration Spokesman* confers an aura of intimate access and sophistication upon the source which is hard to escape (be it news secretary, just secretary, or even chambermaid). Conversely,

one of the many ways of 'belittling' sources is to make the propositional nature of their assertions explicit. To report that *The Russians claim that the Communist Summit Conference was a demonstration of unity within the world of socialist societies* is to cast the shadow of doubt upon what they are claiming. Propositional sentences are most compelling when their propositional status remains implicit.

Students of the Oxford School of Philosophy, notably Wittgenstein and Ryle, have suggested that words can only have meaning when used in sentences. But surely this is a two-way street. Just as sentences constrain the meanings of words, so do the meanings of words constrain the sentences that can be produced in ordinary language.

At the grossest level there are syntactic rules: correctly speaking English, I cannot create sentences of words coded grammatically as Object-Subject-Verb in that order--for example, *Garlic I taste*--although ordinary speakers often do produce just such non-sentences when highly motivated by the logical object (*garlic*). At the finest level there are strictly semantic rules: it is semantically anomalous to say *She plead with him tolerantly*, because *plead with* is coded for a Subordinate relation of *she* to *him* whereas *tolerantly* implies a Superordinate relation of *she* to *him*. This is the level at which language can most subtly constrain how we think about things, and some semantic juggling is often necessary to make true propositions. I think it would have been accurate to say in January 1973 that *Secretary Henry Kissinger plead tolerantly with President Thieu of South Vietnam to accept the terms of the Paris Accords*--tolerantly because of relative power status of the nations involved. There are, finally, pragmatic rules of usage--rules which reflect what we know, or think we know, about the real world. Thus the sentence *He shot at the car with a Reagan sticker* is unambiguous only because we know that you can't use bumper stickers to shoot with.

Why do I claim that sentences are potentially radical? For one thing, whereas the words available to us at any time are finite, the number of sentences we can create with them is potentially infinite. For another thing, most sentences we encounter are novel as wholes; I know that most of the sentences in this paper were novel to me as I wrote them, and I assume they are novel to you, yet I have reason to hope that we both understand them. Furthermore, sentences can be whimsically nonsensical or deliberately nonfactual--we can lie with sentences. But most importantly, sentences 'do things' to the words they use, forcing new and sometimes revealing interpretations of reality. When, of a certain woman, I say *She will make someone a nice husband*--or, so that Woman's Lib. won't come after me, *He will make someone a nice wife!*--I am breaking semantic rules, to be sure, but I am also offering a thumbnail personality sketch of the lady--or gentleman--in question.

Sources and Effects of Sentences

Now let's return to the world of international politics. Precisely what sentences about this world are we likely to create and how are people likely to interpret them? The actual sentences about nations and their relations that we hear are a vary small subset of the possible sentences

which could be produced, even within the rules of sentencing just outlined. The reason is simply that we usually say 'what is in our minds'--but this will require a bit of elaboration. In my analysis of why we create the sentences we do and interpret them the way we do, I shall emphasize only one semantic feature--the evaluative or attitudinal feature. But this aspect of meaning certainly carries the most weight in politics.

Try to imagine a space of some unknown number of dimensions within which the concepts we talk about appear as points. The dimensions of this space are the semantic features already discussed. The most prominent dimension of this space--let's say the up and down of it--is the evaluative, or good-bad, dimension. This dimension of meaning is based upon the human bedrock of pleasures and pains, of rewards and punishments, of threats and promises.

One's own Ego, or selfconcept, is good by virtue of biological imperative--at least until one is taught about sin or becomes old and cynical--and this Ego early-on becomes the arbiter for evaluating other concepts. Signs of immediate things like PAIN, DIRT, and DANGER, which hurt, shame, and threaten Ego, tend to move downward toward maximal Badness. These concepts in turn become pivotal evaluators, pushing and shoving other, more remote things, like FRIEND vs. STRANGER and GOD vs. DEVIL, into appropriate positions along this evaluative dimension--simply by virtue of being associated, positively or negatively, with them in propositional sentences. Exposed to assertions like *My Daddy says Blacks are lazy* or *My teacher says policemen protect us from thieves*, race concepts, occupational concepts, nation concepts, and all the rest gradually acquire values. The underlying principle in this primitive mapping of semantic values onto things is pressure toward cognitive consonance (or, if one prefers the negative, avoidance of cognitive dissonance). It is by such a basic process that the Unfamiliar is assimilated to the Familiar and thereby given meaning. The Cold War is steadily fired by application of its own tired rhetoric, which regularly appeals to the pivotal concepts of GOOD and EVIL for its sustenance--thus being a kind of Holy War.

Assertions connect topics with commentaries. The connectors (verbs) may be either associative (A *is* B, A *likes* B, A *helps* B) or dissociative (A *is not* B, A *dislikes* B, A *hinders* B). If both topic and commentary have the same evaluative sign (both good or both bad), associative assertions will be consonant and dissociative assertions dissonant. If both Russia and China are held to be communist and therefore evil, it is cognitively consonant that *Russia should support China*--and needless to say, from the State Department on down, many Americans found the late President de Gaulle's associative moves toward both Russia and China back in the mid-1960's most unsettling. It is also unsettling for some folks now when there is a developing détente in the relations of the U.S. and mainland China.

If topic and commentary have opposed signs (one is good and the other bad), then dissociative assertions become consonant and associative ones dissonant. In human conflict situations it is easy to believe aggressive dissociative assertions made by one's opponent (*We will bury you*) and

hard to believe his conciliatory associative assertions (*We can coexist peacefully*). It just simply is congruent in Black-and-White thinking for BAD GUYS to threaten GOOD GUYS, and vice versa. This creates a constant bias in *credibility* which encourages escalation of conflicts.

Most people, most of the time, create sentences that are consonant with their own systems of attitudes and beliefs. This is one reason why they produce only a small subset of possible sentences about the world. If someone believes that Communism is evil, he is not likely to produce sentences like *I favor the Communist system* or like *Communists can be trusted*, and so forth for myriads of other potential sentences. But what happens when a person is constrained to produce assertions that are inconsistent with his own attitudes and beliefs? What happens when he is exposed to assertions made by others that are inconsistent with his own system?

This is where what I call psycho-logic enters the picture. This is an entirely illogical, but very potent psychological, process. It is designed to restore the comforting state of mental equilibrium--with all the Goods and Bads in their right places--whenever this state has been disrupted by incongruity. Incongruity creates cognitive stress, and the victim is driven to do something about it; when he does something to the world or to himself which restores cognitive balance, it gratifies him and makes him feel secure. Psycho-logic may not follow logical principles, but it is nevertheless peculiarly significant for an understanding of the dynamics of international politics.

When a person is constrained by his own selfinterest to talk and act as if he agrees with certain people and believes certain things, his behavior is inconsistent with his own values and he is under constant cognitive stress. One resolution is simply to tolerate the dissonance, rationalizing it as leading to a larger goal--but people differ in their tolerance of ambiguity. One wonders if the late Mr. Adlai Stevenson suffered just such prolonged cognitive stress while functioning as U.S. Ambassador to the United Nations. Another resolution is to change one's behavior, get out of the situation, say what one really thinks; people in public life do commit suicide sometimes, politically as well as physically. The most likely resolution, however, is a form of psycho-logic: the person gradually and unconsciously changes his values and beliefs toward consistency with what he feels he must say and do. Men on the accession routes to power are peculiarly susceptible to such self-induced 'brain-washing'.

It can be hypothesized that the same dynamics influence relations among nations. If the people of nations A and B really dislike and distrust each other, but they are constrained by their mutual selfinterest to repeatedly behave as if they liked and trusted each other, their mutual perceptions should become more consistent with their mutual behaviors. Is there any evidence in recent history for such processes at the international level?

In a significant paper titled 'The Kennedy Experiment', Amitai Etzioni documents a real-world test of this hypothesis. Beginning with the late

President Kennedy's speech at the American University on June 10, 1963, in which he announced the first unilateral initiative (that the U.S. was stopping all nuclear tests in the atmosphere), a series of reciprocal unilateral steps were taken (the Soviets stopping production of strategic bombers, the Americans approving the sale of $250 million worth of wheat to the Russians, and so forth), culminating with conclusion of the Test Ban Treaty--clearly to the selfinterest of both sides. Did the predicted psychological side-effects occur? A correspondent to the New York Times had this to say on June 16, near the beginning of the experiment: '. . . there was a new threat of international peace in the air this week, the kind of threat that leaves sophisticates smirking and the rest of us just dumbfounded'; but on September 22 the same correspondent had changed his tune: 'We have cleared the air and cleared the atmospheres and warmed the climate and calmed the winds.' The Kennedy Experiment came to an abrupt end--in Dallas, Texas.

When a person is exposed to assertions from others which cause him cognitive stress, psycho-logic offers a number of defensive mechanisms. He may simply deny the assertions and stop thinking about them. More subtly, he may accept the assertions intellectually but deny their emotional implications; many Americans became quite callous about the use of napalm, crop-destroying sprays, and other indescriminate weapons against living things in South Vietnam. For insurance against cognitive stress, one may selectively expose himself to information that is consistent with his own belief system and selectively avoid information that is inconsistent.

By attributing different motives to ourselves and our opponents--benevolent motives to WE and malevolent motives to THEY--exactly the same behavior by both can be rationalized into the value system. Psycho-logic thus creates a double standard of national morality. During the Cuban missile crisis, the American government defined the weapons being implanted in Cuban soil as 'offensive' in nature, whereas essentially similar American weapons in West Germany, Italy, and Turkey were defined as 'defensive' --and the Soviets, of course, applied exactly the same definitions, but in reverse. After the signing of the Paris Accords, both the South and the North Vietnamese flagrantly violated its provisions--yet, when the WE's did it, it was practical military realism, but when the THEY's did it, it was either crass deceitfulness or naked aggression.

Once unleashed, sentences fly about willy-nilly in all directions. This is praticularly true with modern electronic communications and it poses particular problems for political messages. From the viewpoint of the political speaker, he is always dealing with multiple audiences which have multiple interests and multiple relations to him. When a head of state makes a speech, he can be sure that his own party, the opposition party, labor union members and industrialists, his own military and the opponent's military, to say nothing of general publics both at home and abroad, are all listening in, and what falls harmoniously on the ears of one is bound to fall gratingly on the ears of another. Occasionally one may be able to use a private code for the benefit of some segment of his audience. Ex-Secretary Rusk used to refer to the capital of Mainland China not as Peking but as *Peiping*, the name favored by the Nationalists and therefore an insult to the Communists--of which the American public was

quite unaware. From the viewpoint of the political listener, there is a continuous bombardment of conflicting assertions. Within the U.S., one expects this, and indeed it is one index of democracy. But since the ordinary citizen, no matter where he may be, tends to identify other nations as monolithic personalities, conflicting voices from abroad become most confusing. When the hard-line assertions of a militarist are interlaced with the soft-line assertions of a U.N. diplomat, and both are attributed to the same national entity, that nation obviously is guilty of deliberate doubletalk.

Now, it may be argued that human individuals of even ordinary intelligence, although susceptible to psycho-logic, certainly are more sophisticated in their thinking than this analysis implies. This is quite correct. But when all of the complex interplays among the individuals who make up nations have worked themselves into the amalgam of national policy, what appears in the dialogue between nations seems to represent something close to the lowest common denominator of its individual ingradients. But even for a sophisticated individual, the remoteness of the events talked about and his nearly complete dependence upon the mass media render him peculiarly susceptible to psycho-logic when he thinks and talks about international affairs.

TOWARD A SEMANTIC REVOLUTION

Rules are made to be broken. It is only through breaking them when necessary that we can force our language into a more faithful portrayal of things as they are. When a young lad, accused of a felony, exclaims, 'Ah ain't nevah done nothin' to nobody nohow!', he is guilty of a quintuple negative at the very least--but his claim to honorable character is being vividly made. Winston Churchill's coinage of the phrase *Iron Curtain* broke pragmatic, if not semantic, rules of English, but it certainly provided an apt characterization of the Cold War situation at the time. We need a more lively language to revitalize the stale rhetoric of international politics.

But if rules are made to be broken, then there must also be rules for breaking rules--otherwise we have not revolution but chaos. 'When I use a word,' Humpty Dumpty said, in rather a scornful tone, 'it means just what I choose it to mean, neither more nor less.' 'The question is,' said Alice, 'whether you can make words mean so many different things.' 'The question is,' said Humpty Dumpty, 'which is to be master, that's all.' There are grains of both truth and untruth here. Being master of one's words is not synonymous with being arbitrary in one's use of them. Adaptation of a language to the rapidly changing world is essential, but so is continuity of a language with its past.

As to novelty or innovation in language, a continuum can be traced from monotony, through coinage and metaphor, to verbal magic, and ultimately into chaos. The stale rhetoric of international politics approaches the nadir of novelty; words and phrases predict each other redundantly, semantic features are all in rapport, and little new information is communicated by the parading sentences. If there is too much unpredictability --if too many semantic features are in conflict--a pinnacle of novelty may

be approached, but again little is communicated because of sheer chaos. Just as there is an optimum degree of unpredictability in the arts--variations on themes which maximize aesthetic pleasure--so is there an optimum degree of rule-breaking in language which maximizes communication of fresh ideas, whether it be in poetry or politics.

Effective metaphors break semantic rules optimally. In the literal use of English, one cannot say *the panther shouted* (only humans can *shout*). Speaking poetically, however, one might say *the thunder shouted down the mountainside*, thereby humanizing the thunder; but I, at least, could not say *the brook shouted down the mountainside*--without making it a torrent. *Thunder* and *shouting* share enough features to allow overriding of the Humanness distinction, but *brook* and *shouting* do not--*the brook chuckled down the mountainside* would be fine. I use the term 'verbal magic' to refer to the case where a false context is used to support a semantically anomalous assertion. My favorite example is a TV beer advertisement: After dropping bottles of the brew from skyscrapers, running them over with steamrollers, and flinging them against brick walls--with nary a scratch to the glass bottles--the assertion is brightly made that *this beer has indestructible flavor!*

One ingredient of a semantic revolution is a healthy suspicion of familiar words and phrases. By way of example, let us analyse the familiar phrase, NUCLEAR DETERRENCE. The assertion *A deters B* carries the implications that B must be aggressive (since he requires deterring) and that A is both nonaggressive (since he merely deters) and actually benevolent (since he is opposed to a potential aggressor). The phrase MUTUAL NUCLEAR DETERRENCE, demanded by the real-world fact of Soviet nuclear capability, is a bit more sophisticated--for one thing, it neutralizes A and B with respect to just who is benevolent and who is potentially aggressive--but this phrase also has subtle connotations. It has a stable, reassuring feel to it--almost like being in a medieval suit of armor--but given its foundations on the shifting psychological sands of mutual fear and distrust, nothing could be much less stable or reassuring. In other words, the semantic implications of this phrase simply do not fit its referent.

Another ingredient of a semantic revolution is imaginative flexibility with, and tolerance for, unfamiliar words and phrases. Returning to the concept of MUTUAL NUCLEAR DETERRENCE, with its illusory assurance of stability, it is refreshing to note that one well-known strategist dubbed it 'the delicate balance of terror'. In coining this phrase our strategist was deliberately violating semantic proprieties; *terror* is certainly not something our language implies can be delicately balanced, and yet this is precisely the state of affairs, as the Cuban missile crisis so amply demonstrated. Creating radical phrases is not restricted to academic types, of course. When asked by CBS News on February 20, 1967, to explain the latest military escalation in Vietnam, an Administration spokesman said it represented a 'calculated outburst of impatience' on the part of President Johnson!

New words and phrases rush into the spaces left by the mismatch of language to reality. The greater the rate of culture change the greater the pressure on speakers to innovate. The nuclear age is creating many

empty spaces in our semantic maps of the world, but linguistic innovation is more apparent in its technological than its political aspects. In The Doomsday Dictionary, we find that a *beach* is a unit of fission energy equaling 3000 billion tons of dynamite, or enough to kill half the earth's population by fallout (a term derived from the motion picture, On the Beach), and that a *kahn* is the more 'modest' quantity of fission energy required to liquidate the entire population of one major country without any shelters--300 *kahn* equals one *beach*.

The term *kahn*, of course, is taken from the name of one of the most provocative, innovative--and macabre--writers about strategy in the nuclear age, Herman Kahn. His several books are replete with lively concepts like *catalytic war* (conflict between two major nuclear powers deliberately started by a third power), *escalation ladder* (steps ranging from subcrisis disagreements to all-out nuclear war and its aftermath), and *stark deterrence* (capacity of *overkill* by a factor of 10 or more, so that miscalculation or wishful thinking by an opponent becomes most unlikely). Of course, there are apposite coinages, which highlight the truly significant properties of things, and inapposite ones, which obscure them. The term *overkill* is not particularly apposite, suggesting degrees of killing humans when humans can only die once; it is actually used to refer to the extra intensities of attack required to liquidate a defended target as compared with an undefended one.

Another ingredient of a semantic revolution is respect for the distinction between words and things. This, of course, is the central theme of General Semantics. Actions should speak louder than words, but in international politics actions are often ambiguous as to interpretation. They may be expressive as well as instrumental. The assassination of John F. Kennedy was interpreted variously as a communist, a military-industrial, a right-wing conspiracy, or a CIA plot--depending on where the observer stood politically himself. The conflict in South Vietnam, at the time of U.S. escalation in February 1965, was officially interpreted as 'aggression from the North' but unofficially interpreted by many unattached experts as 'civil war'. Nevertheless, we would be wise to pay more attention to what nations do and less attention to what they say. As long as the feet of Chinese soldiers and their weapons stayed on Chinese soil, for example, we could afford to shrug off their blustering words.

Yet another ingredient is substituting empathy for projection in the effects of our words upon others. Since most of us do most of our talking to people in a community which shares a common set of meanings and values, we have usually been successful in predicting their reactions to our words simply by projecting on to them how we would react ourselves. When we extend this normal process to people of other cultures using other languages, however, we inevitably get into trouble. Projection made it easy for Americans to assume that the South Vietnamese people saw the conflict there as we did--as part of a massive confrontation between the Free World and Communism. However, in terms of Vietnamese experience it was probably, first and most immediately, a struggle for sheer survival in a war they hardly comprehended and second, but more remotely, a struggle for national independence from colonialism. Developing empathy for others

requires exposing oneself to their conditions and experiences--directly if possible but mediately through well selected descriptions if not--and then role-playing the effects of our messages while standing in their mental shoes.

When we label other people's needs, the words we choose may reflect how much projection or empathy is involved. A former official of the Agency for International Development, Byron Johnson, called the acronym AID a hypocritical misnomer, charging that American 'aid' consisted mostly of self-promoting, arrogantly-administered loans that offended the pride of the people in other nations. Despite the benevolent intent, it is depressing how often AID has been given or taken away for military and political reasons, how often it has foisted marginally useful goods on bewildered people in order to relieve gorged American markets, how often it ended up in the pockets of the relatively well-to-do in foreign countries, thereby actually widening the gulf between Have and Have-not people within the countries being 'aided'. The very term *underdeveloped* is an insult to people whose written histories may go back much further than our own, for it equates material, technological progress with cultural development in general. We would be wiser to speak about unevenly developed countries. This phrase at least admits of the possibility that we Westerners are also underdeveloped in some respects, as well as definitely overdeveloped in others--to the extent that Mankind could perish from it.

Another revolutionary phrase I would recommend is *benevolent subversion*: there is no question but what the West--and particularly the United States in the past few decades--has been the most effective force in history for subverting the governments, the economies, and the whole ways of life of other peoples throughout the world. And there is no question but what we have often been benevolent in intent, at least as we see ourselves. But when forced into anomalous confrontation, the words *benevolent* and *subversion* bring into mind obscure but significant features of each--the potential selfishness of benevolence and the potential altruism of subversion.

Perhaps the most important ingredient in a semantic revolution is using radical sentences to compensate for conservative words. Although new words can be coined and old words can change their meanings, they remain essentially conservative and inept for the fresh purposes of the moment. Sentences can crunch words together into phrases which highlight new features and eradicate old; they can formulate assertions about the world which question tired assumptions and suggest new truths. Of course, the phrases may be bizarre and the assertions incredible upon close inspection, but the potential for fresh ways of talking, and hence thinking, is there.

For twenty years I have been espousing a general strategy for international behavior modeled upon what I think I know about interpersonal behavior--a strategy which seems appropriate for a nuclear age. Its formal and descriptively adequate name is Graduated and Reciprocated Initiatives in Tension-reduction. But early-on I discovered that very few people, including myself, could easily remember this complex nominal compound correctly. While doodling at a conference one day I came up with

the acronym, GRIT--which not only provided cues for the salient features of the strategy but also seemed quite appropriate, since grit is exactly what it requires.

This strategy is a form of calculated de-escalation of conflicts: a nation caught in a conflict spiral deliberately initiates sequences of small, tension-reducing moves, these moves being well within its margin of security and being designed to elicit reciprocating steps from the opponent. To the extent that reciprocation is obtained, somewhat larger, more significant steps can be taken, and both parties move cautiously toward a political rather than a military resolution.

My original paper on this (Osgood 1962), appeared in a book sponsored by a group of Democratic congressmen, titled *The Liberal Papers*; the Republican National Committee described my contribution as 'surrender on the installment plan--to which, in the present context, I can only say 'touché'! Debating the feasibility of such a novel approach to conflict resolution--at least as far as sovereign nations are concerned--forced me to question certain unquestioned assumptions about national security in this nuclear age, and, in doing so, to create a few radical sentences of my own. They will provide a final illustration of my thesis.

I had to assert that the primary motive behind the threatening behavior of nations toward others is usually fear. The traditional assumption, consistent with the dynamics of psycho-logic, is that one's opponent is always motivated by aggressive impulses. Accepting even the possibility of fear motivation enables one nation to respond more rationally (and less fearfully) to the blustering behavior of another.

I had to assert that there is no real security in military superiority in a nuclear age. The traditional assumption is that a nation is secure only when it is so strong militarily that no other nation, or combination of nations, would dare to attack it. But, of course, the possessor of such superiority becomes the focus of fear for others and thus the stimulus for continuing the nuclear arms race--which in our time is an invitation to commit suicide. I would even go so far as to recommend a strategy of calculated nuclear inferiority: when two (or more) nations can destroy each other ten times over, as far as being a viable society is concerned, a calculated degree of inferiority probably offers more security (by virtue of being less fear-producing) and certainly gives one an advantage in initiating arms controlling and limiting agreements. Whether because of political intent or economic necessity, I think that this was the strategy of the Soviet Union vis-a-vis the United States until the past decade or so.

I had to assert that having an invulnerable nuclear deterrent makes it possible for a nation to take calculated risks, risks designed ultimately to eliminate the need for the deterrent itself. The usual assumption is that such capability is merely a deterrent against others, and the possessors stand frozen in their mutual threat. But the simple fact of deterring others also provides a nation with a margin of security within which to take calculated risks. Rather than being seen as simply a continuation of the military tradition of discovering bigger and better ways of destroying opponents, the grisly fact of nuclear weaponsy could be viewed as the impetus toward eradicating the very tradition that gave rise to it.

I had to assert that in the reduction of international tensions, POST commitment by reciprocation can be substituted for PRIOR commitment by negotiation. Although in our ordinary relations with wives, children, neighbors, and colleagues we almost never negotiate a prior commitment before starting tension-reducing actions ourselves, it has been traditionally assumed that under conflict conditions one nation cannot be decent to another without first obtaining from the other an iron-clad prior commitment that it will be decent in return. Not only can patterns of initiatives and reciprocations (post-commitments) reduce mutual hostilities, they can also create an atmosphere within which serious negotiations on critical issues can be more successfully undertaken. Both of these propositions were demonstrated in The Kennedy Experiment, and there have been some other demonstrations in human history.

And, finally, I had to assert that goodwill among nations is a result of, rather than a prerequisite for, de-escalation of tensions. The traditional assumption has been that a nation cannot risk making even conciliatory gestures toward another unless some degree of mutual goodwill and trust exists between them--since otherwise such gestures are likely to be interpreted as 'signs of weakness'. Quite to the contrary, as I argued earlier, given even a modicum of mutual self-interest, the reciprocal actions taken in its service can literally create mutual goodwill where little or none of it existed before.

Interlocking assertions like these can provide a framework for rethinking our relationships to each other on this shrinking little planet at the dawn of its nuclear age. The assertions may be proven false, of course, but so may the tenets of any theory. The important thing is to use the potential radicalness of our sentences to at least reach for a clearer perspective on our own time and place. Revolutions do not guarantee a better world--they only make it possible.

NOTE

*An earlier version of this talk, with the same title, was published in Social Psychology and Political Behavior, ed. by G. Abcarian and J. W. Soule, 1971, Columbus, Ohio: Merrill.

REFERENCES

Kaplan, D. M. and A. Schwerner. 1963. The domesday dictionary. New York: Simon and Schuster.

Osgood, C. E. 1962. 'Reciprocated initiative.' The liberal papers, ed. by J. Roosevelt. New York: Anchor Books.

THE THREE SCALES OF DIACHRONY[1]

William S-Y. Wang
University of California at Berkeley

As always, it is both an honor and a pleasure to speak at the forum of the Linguistic Institute. In the present case, my pleasure is very much enhanced by the theme of this particular Institute, which is 'Language form and Language function: A Western and Nonwestern Perspective'.

It is all too tempting for us to draw conclusions about what human language is like based on the samples of language we most commonly encounter here in the West.

Let me give just one example of this western bias, one with immediate importance. Under the impact of Western languages, it is often assumed that alphabetic orthographies, constructed on phonetic segments, are the optimal system for representing language. This assumption is mostly made tacitly, to be sure, and thus it may be all the more misleading. On the other hand, orthographies based on other units, found outside the West, such as syllabaries or logographs, are felt in some way to be under-developed or retarded, and doomed to fall in time to the ultimate and inevitable triumph of the alphabet.

But clearly this is an issue that is much too important to be assumed a priori on the skewed sample of just the Western languages. It is encouraging to note, though, that serious work is beginning to be done on the question of optimal orthography, based on considerations of language structure and psychological experiments on reading.[2] Hopefully, an increasing amount of involvement will be forthcoming from linguists in such research.

So the moment is long due that we extend our focus beyond the familiar scene of Western languages in a serious way. In this respect, it is reassuring to look through the catalog of this Linguistic Institute, with its rich array of courses on the languages and linguistics of Africa and Asia. The planners of this Institute deserve our warm thanks for their effort in highlighting a more balanced perspective on the study of human language.

The topic I was assigned to discuss this evening is 'Language Change'. Change involves the differences between two or more states in time. Presumably the goal here would be to reveal the principles according to which these differences are implemented, and to discover the mechanisms which brought about these differences. I hope to make some observations on both of these issues of implementation and actuation. But first it is necessary to distinguish change along several time scales, for it seems that the questions, the data, and the methods would not be the same for all these scales.

MICROHISTORY

For convenience of exposition, I will refer to the three time scales as microhistory, mesohistory, and macrohistory. They all deal with change across time and therefore are all aspects of the diachronic study of language. The microhistory of language is reckoned across a very thin slice of time, in years or decades. It is concerned with what Wm. Labov (1972) calls 'change in progress', which offers a diachronic way of looking at synchronic variation.

In the microhistory of language the interests of several research areas converge. The sociolinguist focuses his attention on groups of people, as these are divided by age, by sex, by region, and by social background. Typically, within so short a period we find the model and the copies, the unchanged and the changed, existing side by side in an orderly profusion, vying for survival. By correlating the social parameters with the noted language differences, we may hope to make some short-term predictions on which of the various usages of today will continue into the language of tomorrow (Weinreich et al. 1968).

Closely related to these questions is the study of language contact. With the rapid increase in population and mobility, the notion of a pure and homogeneous speech community, whatever credibility it may have once held, cannot realistically be maintained. Diversity is the fundamental ingredient of change, as is well recognized for biological systems; this dogma is no less true for linguistic systems at every level. The further heterogeneity that contact produces adds yet new dimensions to the challenges of the student of language change.

Among these various groups of people there is a very special class, with a unique biological and social status--and these are the very young. How children learn language--the transmission across generations--is clearly one of the vital questions in the whole of language change. Unlike that of the other groups, the language children use changes not only as a function of differences in environment, but also as a function of an increasing biological capacity due to neural and motor-sensory maturation. Ontogeny hardly ever recapitulates phylogeny with fidelity. Nevertheless, some insights can be gained here that may prove useful toward understanding the processes of language evolution.

Until recently, theorizing on language acquisition, especially within the framework of generative grammar, has been often cast in rather global terms--in terms of whole categories of sounds and the addition and reorganization of rules. However, more fine-grained work done in this area, by Ferguson and Farwell (1975), and Hsieh (1972), and others, clearly show that the real situation is much more complex, both within the case histories of each individual child, and in the strategies and development across different children.

Some children, for instance, are extraordinarily pliable and are eager to experiment with new sounds and new sequences. Others are more conservative in their phonological behavior, and produce only forms that they have a good chance of getting right; they seem to prefer learning

and using those words in which they have phonetic confidence, while avoiding others. Linguistic personalities, it would seem, are manifested at quite an early age, and these differences in strategy may very well have lasting influences on their adult language behavior. There is reason to expect that some of the individual differences in the language of adults can be traced back to the early years. It is an important finding that these differences can be detected even during the learning of the first several consonants, as is shown in some recent work of Ferguson for the acquisition of English laterals, reported in Fillmore, Kempler, and Wang (1979).

From such fine-grained research, it emerges clearly that there is a primacy of lexical learning during phonological development. Even a relatively small scale sampling of the sort done by Ferguson and Farwell shows that the child does not progress by learning units like phonemes or allophones, but rather by gradually adding lexical items to his repertoire. The same sound in similar contexts may undergo altogether different histories, as this sound appears in different words. This is clearly demonstrated in Table 1. These data in the table are extracted from the study of Ferguson and Farwell, who followed with great care the development of three children who were learning their first fifty words. The data are of one child's progression in the acquisition of initial *b*- across a span of some four weeks. They show how different words exhibit different patterns of variation according to different schedules. The unity of the phoneme only emerges at the end of the acquisition process when mastery is complete.

A CHILD'S ACQUISITION OF SOME *b*-INITIAL WORDS IN FOUR WEEKS (BASED ON FERGUSON AND FARWELL 1975)

	Session VI	Session VII	Session VIII	Session IX
baby	b~β	b~w~p^h	b	b
book	b~ɸ	b	b	b
bye-bye	b~p^h	b~β	b	b
ball	b	b	β	b
blanket	b	b	-	b
bounce	b	b	-	-
bang	-	b	-	-
box	-	b	-	b

Table 1

The basic unit of acquisition, then, is something like the word. The awareness, conscious, or subconscious, of phonetic identity or similarity

between portions of words, is something that comes to different children in completely idiosyncratic ways, and probably cannot be attributed to any single uniform stage of development. This theme of lexical primacy is central to my remarks there, and I will return to it later.

MESOHISTORY

As opposed to microhistory, the great bulk of the literature of language change actually deals with the middle time scale. Historical linguistics has traditionally concerned itself with changes that occur across centuries or millenia. Since it reaches further back in time, the primary data for language history are much more uneven and uncertain. Written records of early languages are low in both quality and quantity, and only in a handful of cases do they extend for more than 2 or 3 thousand years. The methods of reconstruction have their intrinsic limitations. Most of us would probably agree with Kiparsky (1976) when he set the outer limits at 10 to 20 thousand years 'over which we can hope to reconstruct anything at all about Proto-language'.

A classic question in language mesohistory, dating back to at least the neogrammarians, has been the manner or means by which a change is implemented. This question has caught the attention of a wide gamut of scholars through the decades, from Henry Sweet to Alf Sommarfelt and Sapir, to Hoenigswald and Halle.

The received doctrine on this question has long been one which may be characterized as phonetically gradual and lexically abrupt. The idea is, that once a phoneme changes, it changes in all the relevant words according to the same schedule. By the claim of lexical abruptness, the regularity of the change, the so-called neogrammarian hypothesis, would follow as a consequence by definition. But since a change may involve hundreds of words, and since languages do not seem to effect such wholesale changes in pronunciation overnight, the phonetic gradualness becomes a necessary corollary of lexical abruptness.

So the notion of lexical abruptness is motivated by believing that changes are always regular, and the notion of phonetic gradualness is necessitated by believing in lexical abruptness. And it is essentially this doctrine of historical change, with minor variations, that has retained the most widespread acceptance in the work in generative phonology as well as the structuralist phonology that preceded it.

But here again, as was the case in the microhistory of language, things are not that simple. Empirical investigations over the past decade or so on a variety of languages, using large quantities of data, have shown that there are changes which are implemented in a manner that is lexically gradual. That is, a change may initiate on a handful of words in the lexicaon, where these words do not constitute any natural phonological or morphosyntactic class. Then the change affects an increasing sector of the lexicon in time, perhaps eventually completing its course on all relevant words. Whether or not a particular change ever completes its course depends on a whole host of factors, some of which are outside the linguistic system. We can get at least a rough idea of

the degree of regularity that sound changes exhibit from a study that C. C. Cheng and I (Cheng and Wang 1971) did on the development of initial consonants in Chinese over the past 1400 years, shown in Table 2.

Time and again, scholars of sound change have observed that exceptions to correspondences are embarassingly numerous, frequently the irregular reflexes outnumbering the regular ones. Table 2 gives a rough quantitative index for these observations. Each cell in the table refers to the development of one Middle Chinese initial consonant (ca. 600 A.D.) into one of the contemporary dialects. When the development is indeed perfectly regular, the cell is marked with an 'x'; otherwise it is left blank. Each blank cell in Table 2, then, is one of the 'embarassingly numerous' cases where regularity was not reached--and there are indeed many of them.

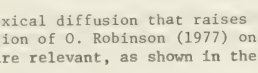
A particular change, for various reasons, may even reverse its course. Tore Janson (1977) has studied such a case in the deletion of final *-d* in the Swedish of Stockholm, where he attributes the reversal to the influence of orthography. At any rate, this picture of implementation which involves a gradual change across the lexicon, which may or may not result in complete regularity, has been called lexical diffusion (Wang 1969).

Since lexical diffusion is a process via which a change is implemented, it is in principle accessible to any change, however the change is actuated, i.e. whether the actuation is external or internal to the system under change, or whether it is phonetically or analogically triggered, etc.

An interesting case of lexical diffusion has been reported by Lyovin (1977), where the actuation is due to the avoidance of homophony within the inflectional paradigm of Tibetan verbs. The study is of particular theoretical significance because of the intricate relations between a change that reduced initial consonant clusters, which causes homophony, and the hierarchy of verb categories which controls the schedule according to which periphrastic expressions are created to avoid the homophony. The present tense dominates the past tense, for instance, so when the two forms become homophonous due to cluster reduction, it is the past tense form that gets displaced from the inflectional paradigm. Similarly, the past tense form dominates the future tense form, and it is the latter that gets displaced when homophony occurs. In each case, the meaning of the displaced form has to be expressed periphrastically.

Lyovin's examination of the data across the various categories of verbs in Classical Tibetan led him to observe that 'homophony-inducing sound changes will be blocked from applying to certain lexical items or classes of lexical items until compensatory developments permit the further diffusion of the shifts in question' (p. 129).

Another study in lexical diffusion that raises important theoretical issues is the investigation of O. Robinson (1977) on Swiss German vowels. Two historical changes are relevant, as shown in the diagram.



MIDDLE CHINESE INITIALS WITH SINGLE REFLEXES
IN MODERN DIALECTS
(Cheng and Wang 1971)

	b a 35	b b 37	m — 91	v a 13	v b 21	ɱ — 22	t^h — 60	d a 43	ɖ a 27	g a 24	ts^h — 54	dz a 27	s — 99	z b 15	tʃ — 26	$tʃ^h$ — 27	dʒ a 8	dʒ b 12	ʃ — 50	tɕ — 70	$tɕ^h$ — 28	dʑ a 6	dʑ b 10	z a 19	z b 30	ʔ — 86	Total
Beijing	x		x	x	x		x	x	x								x			x	x		x			x	12
Jinan	x		x	x	x		x	x	x								x										8
Xi'an	x		x	x	x			x																			5
Taiyuan			x	x	x	x			x							x	x		x				x				9
Hankou	x		x	x	x	x		x								x	x										8
Chengtu			x	x	x				x							x							x				6
Yangzhou			x	x	x		x									x	x		x							x	8
Suzhou	x	x		x				x	x				x		x	x	x	x	x	x	x	x	x	x	x	x	18
Wenzhou	x	x	x					x																			4
Changsha			x																								1
Shuangfeng			x	x				x																			3
Nanchang			x	x			x	x								x	x		x				x				8
Meixian	x		x							x					x		x			x	x	x	x				9
Guangzhou	x		x	x	x			x		x		x			x	x											9
Xiamen																											0
Chaozhou																										x	1
Fuzhou			x								x			x								x	x	x			6

Table 2

Robinson notes that the 'rounding rule was diachronically much later than umlaut'. Both rules continue to operate synchronically in morpho-phonemic alternations.

Now if the rules applied synchronically in the order in which they entered the language diachronically, then the /a/'s which are subject to the umlaut rule should change to /æ/'s, and thus escape the rounding rule. But in almost all of the northern dialects, these /a/'s actually umlaut to /ɔ̈/'s, presumably via /ɔ/. Furthermore, a synchronic rule that rounds /æ/ into /ɔ̈/ would not be possible; it would be contradicted by other data. For this interesting phenomenon, Robinson adopts Kiparsky's interpretation, that the synchronic rules have reversed their diachronic order, that is, from umlaut-rounding to rounding-umlaut.

The data of relevance here have to do with certain northern dialects, e.g. Kesswil, where many forms exhibit the morphophonemic alternation /a~æ/, when the reordering hypothesis would expect /a~ɔ̈/. Robinson's explanation of these æ-alternations is that rule reordering, though an abstract change, is diffusing across the lexicon, much as concrete changes like *d*-deletion or vowel shift. These æ-alternations, then, are forms which the reordering has not reached as yet. The theoretical significance of this explanation is the extension of the process of lexical diffusion to operate at a relatively abstract level of sound change. To support his explanation, Robinson gives some exemplary arguments against some alternative solutions, including the possibility of borrowing these æ-alternations from other dialects.

Work on lexical diffusion over the past decade has shown that the traditional view of sound change has been excessively restrictive, and that the implementation may very well proceed along other paths than one which is lexically abrupt. This work has been surveyed in Chen and Wang (1975) and Wang (1976, 1979), and partly anthologized in Wang (1977). At this stage, there is less need to document additional cases of lexical diffusion. Rather, our next challenge, it seems to me, is to solve the puzzle of what kind of sound change would travel along which path for its implementation. In addition to the lexical and phonetic parameters, there is the additional statistical parameter that must also be considered in investigating the many paths of sound change.

Two other directions of current research in this area should also be mentioned, in view of the very promising results obtained so far. One is the demonstration of the relationship of the relative frequency of words to their schedules of change, as discussed by Hooper (1977) for English, and Gerritsen and Jansen (1978) for Dutch. Although such a relationship has been hypothesized at least since a century ago, it is only recently that it can be objectively tested as a result of the availability of frequency dictionaries. It is of special importance that Hooper has been able to show that high frequency words in English are leaders in a phonetic change like schwa-deletion, but are laggers in an analogic change like the weakening of strong verbs. If we can continue to identify independent variables for the chronological profiles of change (relative frequency being one such variable), then perhaps the goal of being able to make short-term predictions in sound change will be eventually attainable, after all.[3]

The other direction that I'll briefly allude to has to do with subgrouping languages. Since it has been assumed in the bulk of historical work that sound changes are all lexically abrupt, it follows that for any change in question, a particular group of languages either shares the change or not--a binary yardstick. But once we recognize that diffusion processes may be at work, then some changes at least can be exploited for much more information. Indeed, as Hsieh pointed out in his seminal article (1973), this was the insight underlying Swadesh's method of lexicostatistics. Instead of a binary yardstick, quantitative measures can be made on various subgroupings of the languages on the basis of both changed and unchanged cognates. Hsieh demonstrated the effectiveness of this approach on the Wu dialects of eastern China by using a single tone change.

More recently, Bh. Krishnamurti (1978) and Krishnamurti and his colleagues, in some yet unpublished work, have made important progress in developing a suitable methodology for the subgrouping problem. Using the Dravidian Etymological Dictionary as the starting point, they have analyzed a large amount of computerized data to determine the genetic relationships among various languages of South India. There is every indication that their procedures will prove fruitful when validated against linguistic situations where the answers are known, at which point they can be used in situations where no answers are available (for unwritten languages, for example).

MACROHISTORY

In considering language change within the largest time perspective, its macrohistory, we are faced with the greatest challenges--since the relevant primary data are mostly not available. Furthermore, the necessary backgrounds and methods are frequently outside of the typical domain of linguistics. So we need to look into neighboring disciplines that work on related questions, such as anthropology, biology, ethology, psychology, zoology, and so on. An example of the interdisciplinary convergence of interest in this area is the very successful conference in the origin and evolution of language and speech, sponsored by the N.Y. Academy of Sciences in 1975 (Harnad et al 1976). See the excellent discussion of this conference by Hill and Most (1978).

A useful assumption to make here, one which most investigators accept, is that language did not abruptly burst into the course of human evolution fully in the state of intricacy and complexity that we find it today. To believe the so-called discontinuity theory, that there was some sudden wholesale genetic mutation that in a fell swoop transformed muteness into eloquence is to relegate macrolinguistic research to the mystical regions of miracles and spontaneous creations.[4]

Rather, there must have been a very long course of emergence, much longer than the span of time over which current methods of historical linguistics can take us. During this emergent state, protolanguage probably at different times shared various developments, by either homology or analogy, with the communication systems of other species. So it is likely that sorting out the ingredients of these other systems will

contribute to our understanding of how the ingredients of human language were accumulated piece by piece over the hundreds of millenia.

One category of ingredients that has been present since the early stages of emergence is the use of gestures. Although gestures are used across a wide range of species, we share many specific ones with other primates, including several facial expressions, presumably because of the shared anatomy. Gestures accompany speech in all languages, time-locked to various extents; they also play a relatively more prominent role in the early communications of very young children. Phylogenetically, Gordon Hewes (1973), among others, has championed the view that human language originated in gestures.

Another category of ingredients that is also widespread in communication systems is the use of prosodic features--fundamental frequency, duration, and intensity. These features are probably exploited more fully in bird song than anywhere else, though recent investigations on primate calls are revealing a variety of new ingredients, such as amplitude modulation, frequency modulation, and the use of formants. In human evolution, there was a correlated development of the increased use of the hands for fighting and carrying with the increased use of the mouth for communicating. There are certain obvious advantages that prosodic features have over gestures: Communication can take place over larger territories requiring no visual contact (such as in darkness or across foliage), while the rest of the body can be engaged in other simultaneous activities (such as running or fighting). So it is reasonable to expect that the prosodic communication gradually expanded in importance relative to the gestures.

As the civilization of early man became increasingly complex, however, he needed larger sets of signals with which to communicate the expanding vocabulary of messages. Since communication facilitates the planning and execution of group activities, for purposes such as hunting for food or defense against predators or other tribes, we can easily see how groups that have developed better language would have better chances of survival. Everything else being equal, better language enables the formation of larger (and hence stronger) groups, the transmission of more precise and varied information across both space and time, and the perpetuation of the gene pool of the group of users of that language. Biologically speaking, then, selectional pressures favored those populations with better tools for communication in probably much the same way that they favored better tools for digging or fighting.[5]

As the message set grew, the three prosodic parameters became insufficient as carriers for the signals. A phonology that was largely prosodic (and laryngeal) eventually changed into a segmental one, making use of the richer possibilities of the supralaryngeal gestures. Whereas other primates share with us the use of gestures and prosody, it is the transition from prosody to segmentals that is the critical step which our ancestors alone have taken. It has been proposed by P. Lieberman (1975) that this transition was made possible by the phylogenetic descent of the larynx, and that the descent was a special adaptation for speech. A more plausible view to my mind is that the descent of the larnyx, as well as a whole host

of other skeletal restructurings in the head and neck, was one of the many mechanical responses that the human body had to make in assuming erect posture. In this latter view, where the structure was largely available before the linguistic use, it would be more accurate to call it a case of preadaptation.

In any case, segmental phonology grew dominant over prosody, just as prosody in an earlier stage of emergence dominated gestures. All three systems continue to co-exist in every language today, functioning in mutually supplementary ways. The importance of the transition into segmental phonology is highlighted by recent researches on communicating with apes. Whereas chimpanzees show remarkable cognitive abilities at symbolization, including the processing of complex sentence types, they have next to zero ability at producing controlled segmental sounds. From an evolutionary perspective, that critical step which early man took in pairing messages with the particular medium of consonants and vowels was the start of the journey that led to speech, the 'indispensible foundation' upon which language is built. Judging from the amount of new brain tissue in the cortex that appears to be involved with language, that step must have been taken quite some time ago (cf. A. Liberman 1974).

Although there may never be any way of documenting the transition directly, we can imagine how it could have come about by referring to processes of phonological change that can be observed in recent times. These are of the type that has been called 'phonologization'. Typically what happens is as follows. At time t_1, one group of words is distinguished from another group in that the former group has the phoneme X while the latter has the phoneme Y. For reasons of coarticulation, X and Y act upon their environments in different ways, i.e. XE' but YE", even though at t_1 the difference between E' and E" is not considered distinctive, and may not even be noticed. But should a merger take place by the time t_2, when X > Y, then the two groups of words are no longer distinguished by X and Y, but now by E' and E". The difference between E' and E" becomes 'phonologized' at t_2--it has become phonemic.

There are numerous such cases in the literature. Chinese, for one example, had distinct voiced and unvoiced stops in syllable-initial position. Because of coarticulation, the syllables with voiced stops must have had a different pitch contour from those with unvoiced stops. (Compare the different pitch contours for the English words *bin* and *pin*.) Some centuries ago, however, the voiced stops merged into the unvoiced ones. Now the two groups of words are kept apart no longer by the voicing in the initial stop, but by the tonal difference that was once caused by that voicing.

The transition from prosodic utterances to segmental utterances could not have been this straightforward, of course, and must have taken place over a much longer timespan, probably with many false starts in the process. The segments started as conditioned or spurious accompaniments of the prosodies, and were only eventually 'phonologized' by transference. Once the pairing between messages and the segmentals got established, however, language must have evolved at an explosive pace, given the much greater signalling potential and physiological economy of a segmental phonology over alternative media.

Over the past several years there has been some significant research into the sign language of the deaf. This work is of fundamental importance for the light it sheds on language viewed from a different modality, and thereby giving us a broader perspective on human communication. Experiments have been done on comparing the rates of production and perception of speech with those of sign language. While such experiments are of great intrinsic interest, it should be clear that certain questions of linguistic macrohistory cannot be directly answered from their results. For even if it can be completely demonstrated that sign language is as effective as speech in every conceivable way, which I would find totally surprising, the point remains that the user of the sign language is the direct beneficiary of numerous millenia of language evolution and elaboration that had in fact taken place via speech. The deaf is deprived of only a minor, peripheral part of that total neural machinery. The equivalence between modalities, if it can be demonstrated, tells us that once the full symbolization system has been developed in our brain, a surrogate medium may serve as well as speech. It tells us nothing on whether the surrogate medium had enough to recommend it for language to have evolved into its present state of intricacy and richness, which is clearly a stronger requirement.

In contrasting the three scales of diachrony, I have perhaps put too much emphasis on the differences between them. Clearly the three must be related to each other in intimate ways, since they are but three different time windows through which we are viewing the same phenomenon. At present, the facts are scanty. There are huge gaps in our knowledge on the steps in which language evolved through the successive emergent states into the present steady state. It is obvious that the study of linguistic universals must be centrally relevant here in providing a base line for the steady state. Similarly, our understanding of the relation between microhistory and the classical concepts of linguistic change is spotty at best. A great deal of basic research needs to be done before a plausible scenario can be provided for what has been called 'the life and growth of language'.

NOTES

[1]This paper is prepared while I am a Guggenheim fellow visiting Osmania University in India. Thanks are due both to the Guggenheim Foundation and Bh. Krishnamurti of Osmania University for their support and encouragement. I would also like to acknowledge the many years of camaraderie and collaboration of C. C. Cheng, which were critical ingredients in the progression from idea to data to knowledge on language change.

[2]A good beginning is the Cross-Cultural Conference on Language, Reading and Orthography, sponsored by the National Institute of Child Health and Human Development, Bethesda, Md., September 1978. The proceedings are being edited for publication by J. Kavanagh and R. Venezky. An extensive discussion of Chinese orthography from several viewpoints is available in the Journal of Chinese Linguistics 6.2, 1978; see especially the contribution by Ovid Tzeng et al. for recent psycholinguistic work. (See also Hardych et al. 1977, 1978.)

[3]Of all the independent variables that influence change, probably the least stable and therefore most difficult ones to capture are the social ones. Note the perceptive counsel of a poet on this point, Alexander Pope:

> In words, as fashions, the same rule will hold
> Alike fantastic, if too new or old
> Be not the first by whom the new are tried
> Nor yet the last to lay the old aside.

[4]There was, of course, no dearth of proponents for such discontinuity theories among philosophers working before evolutionary thinking was developed in biology. See Stam (1976) for a useful historical critique, especially from mid-18th century through mid-19th century in Europe.

[5]Once languages have reached a more-or-less steady state, these selectional pressures no longer operate as they did during the emergent state. Labov's observations on language diversification being dysfunctional (1972:273) are presumably based on the steady state changes, as indeed were the remarks by Charles Darwin and Max Muller that Labov referred to.

REFERENCES

Chen, M. and W. S-Y Wang. 1975. Sound change: actuation and implementation. Language 51:255-81.

Cheng, C. C. and W. S-Y. Wang. 1971. Phonological change of Middle Chinese initials. Tsing Hua Journal of Chinese Studies 9, 1. 215-70.

Ferguson, C. A. 1979. Phonology as an individual access system: some data from language acquisition. In Fillmore, Kempler, and Wang.

------ and C. B. Farwell. 1975. Words and sounds in early language acquisition. Language 51. 419-30. Reprinted in Wang 1977, 7-68.

Fillmore, C. J., D. Kempler, and W. S-Y. Wang. 1979. Individual differences in language ability and language behavior. New York: Academic Press

Gerritsen, M. and F. Jansen. 1978. Word frequency and lexical diffusion in dialect borrowing and phonological change. Studies in Dutch Phonology. Amsterdam.

Hardyck, C., O. Tzeng, and W. S-Y. Wang. 1977. Cerebral lateralization effects in visual half-field experiments. Nature 269. 705-7.

------. 1978. Lateralization of function and bilingual decision processes: Is thinking lateralized. Brain and Language 5. 56-71.

Harnad, S. et al. (eds.). 1976. Origin and evolution of language and speech. Annals of the New York Academy of Science 280.

Hewes, G. W. 1973. Primate communication and the gestural origin of language. Current Anthropology 14. 5-24.

Hill, J. H. and R. B. Most. 1978. Review of Origin and Evolution of language and speech, ed. by Harnad et al. Language 54. 647-60.

Hooper, J. 1977. Word frequency in lexical diffusion and the source of morphophonological change. Current progress in historical linguistics, ed. by W. Christie, 95-105. Elsevier: North Holland.

Hsieh, H. I. 1972. Lexical diffusion: evidence from child language

acquisition. Glossa 6. 89-104. Reprinted in Wang 1977, 133-47.
------. 1973. A new method of dialect subgrouping. Journal of Chinese Linguistics 1. 64-92. Reprinted in Wang 1977, 159-97.
Janson, T. 1977. Reversed lexical diffusion and lexical split: loss of -d in Stockholm. In Wang 1977, 252-65.
Kiparsky, P. 1976. Historical linguistics and the origin of language. Annals of the New York Academy of Science 280. 97-103.
Krishnamurti, Bh. 1978. Areal and lexical diffusion of sound change. Language 54. 1-20.
Labov, W. 1972. Sociolinguistic patterns. Philadelphia: University of Pennsylvania Press.
Liberman, A. M. 1974. The specialization of the language hemisphere. The neurosciences: third study program, ed. by F. O. Schmitt and F. G. Worder. Cambridge, Mass.: MIT Press.
Lieberman, P. 1975. The origins of language: an introduction to the evolution of human speech. New York: MacMillan.
Lyovin, A. 1977. Sound change, homophony, and lexical diffusion. In Wang 1977, 120-32.
Robinson, O. W. 1977. Rule reordering and lexical diffusion. In Wang 1977, 69-85.
Stam, J. H. 1976. Inquiries into the origin of language: the fate of a question. New York: Harper and Row.
Tzeng, O., D. Hung, and L. Garro. 1978. Reading the Chinese character: an information processing view. Journal of Chinese Linguistics 6. 287-305.
Wang, W. S-Y. 1969. Competing changes as a cause of residue. Language 45. 9-25.
------. 1976. Language change. Annals of the New York Academy of Science 280. 61-72.
------ (ed.). 1977. The lexicon in phonological change. The Hague: Mouton.
------. 1979. Language change: a lexical perspective. Annual Review of Anthropology 8. 353-71.
Weinreich, U., W. Labov, and M. I. Herzog. 1968. Empirical foundations for a theory of language change. Directions for historical linguistics, ed. by W. Lehmann and Y. Malkiel, 97-195. Austin: University of Texas Press.

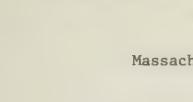
ANALOGICAL CHANGE AS A PROBLEM FOR LINGUISTIC THEORY[1]

Paul Kiparsky
Massachusetts Institute of Technology

1. *Simplification vs. proportions.* Recent linguistics has moved from the structuralist segregation of synchronic and diachronic study of language towards a reintegration of the two, based on the idea that structure can determine change, with the corollary that change can therefore be diagnostic of structure. This program has had consequences on both sides of the equation. On the synchronic side, it is mainly through evidence from change that phonologists have argued for such factors as opacity and recoverability in the evaluation measure, and for the imposition of restrictions on the abstractness of underlying phonological representations. On the historical side, the new conception of linguistic structure evolved through efforts to construct explicit formal generative grammars has provoked some rethinking of the nature of both sound change and analogy.

As an example of the relevance of synchronic work for historical linguistics we may take the classic issue of whether sound change admits nonphonetic conditioning. Formal grammar as well as sociolinguistic research on linguistic variation has shown that phonology is intimately connected to higher linguistic levels. This undermines the structuralist concept of autonomous phonology, and thereby in turn casts doubt on the companion concept of 'blind' sound change. Morphological and other nonphonetic conditioning of sound change was previously ascribed to grammatical analogy secondarily acting upon the results of a quasiphysiological process of sound change. It is now more simply taken to be part and parcel of the phonological process itself, as a reflection of the inherent contextually determined and functionally governed variability characteristic of all new rules. This permits a more restricted and therefore more predictive theory of analogy.

In general, the theory of grammar has opened up the possibility of structural explanations as alternatives to be considered in many cases where so far only a historical explanation was available. Noting a change (A > A', B > B') in two classes of items, the traditional historical linguist would tend to seek its origin in one class and postulate its analogical spread to the other:

(1)


In many cases, however, there turns out to be a structural explanation in the form of a shared property of A and B which allows us to assume a common

origin for the phenomenon in both classes:

(2) A , B
↓
A', B'

Thus the investigation of clusters of co-occurring changes becomes an important method of probing into the organization of linguistic phenomena.

The issue to which I should like to address myself here is analogical change.[2] Traditionally visualized as the extension of surface patterns (in terms of proportional schemata) it has more recently been given another interpretation as the elimination of arbitrary complexity in the linguistic system (in terms of the evaluation measure developed in the theory of grammar). At the back of this lies a new view about the nature of the concrete process of analogical change and how it relates to the acquisition and use of language. A proportional view of analogy fits naturally into a theory of language acquisition based on substitution-in-frames techniques and equivalent 'taxonomic' devices. The idea that analogy is simplification of grammar jibes better with the idea that language acquisition is based on a general rule schematism in conjunction with an evaluation measure which selects the simplest grammar from the set of alternatives compatible with the data encountered by the learner, and assumes that learners progress to the adult grammar through a series of intermediate stages of increasing adequacy and complexity. Analogical changes are then interpretable as retentions of features of these intermediate stages into the adult system.

The difference between the views is not so radical in practice as it seems from the above formulation. Both approaches, to be at all plausible, require some elaboration, which may take and in fact has taken different terms. Other factors than simplification have been introduced by proponents of grammar-based theories. An example is transparency (Kiparsky 1971, 1973), which determines the directionality of rule reordering, indifferent as far as simplicity is concerned. Proposals have been made to bring in paradigmatic factors independently of simplicity considerations (Harris 1973; Kiparsky 1972). The general effect of these modifications is to bring the theory into closer agreement with proportional theories. Under the radically 'concrete' analyses of Natural Generative Phonology and Upside-Down Phonology the treatment of analogy as simplification yields predictions which resemble those of proportional theories.

On the other hand, diverse proposals exist to replace bare proportions by interpreted ones. Typical is this passage from von Wartburg (1969:59): '. . . analogy does not operate blindly. The Neogrammarians, and Saussure also, made the mistake of regarding analogy as a purely proportional process. Paul (p. 116) simply asserts that, for a new form to arise, there must be three elements present in the equation. He cites as an example *animus* : *animi* = *senatus* : *x*; x = *senati* (instead of Class. Lat. *senatus*). This conception accounts for only a limited number of instances of the working of analogy. In most cases, analogy does not arise merely as the result of a formal resemblance; the two words so linked are also closely linked by their semantic environment.' Similar ideas are found earlier in Hermann (1931). Another kind of 'annotated proportion' was developed by

Kuryłowicz (1956, 1964, and many other works). Kuryłowicz argued that both the terms of the proportions and the relation between the terms have to be interpreted in terms of certain abstract structural notions. For example, the relation between *a* and *b* and between *c* and *x* in the proportional schema could not be just any relation, but only that between a 'forme de fondation' and a 'forme fondée'. And the terms themselves may have to be different from what is actually pronounced in the language, as in the following proportions:

(3) a. $\underset{o}{R}$: eR = e : X (X = ee = $\bar{e}$)

b. ReT : RuT = RewT : X (X = RuwT = $R\bar{u}T$)

(3a) was proposed to account for the rise of *vṛddhi* (long grade) in nouns (Kuryłowicz 1968:298). (3b) was proposed as the explanation of the Germanic verbs of the type *lūkan* (Kuryłowicz 1970). These proportions critically depend on the equivalence within the early Indo-European phonological systems of syllabic and nonsyllabic resonants ($R = \underset{o}{R}$) and long vowels and geminate vowels ($\bar{e}$ = *ee*).

In this way, analogical proportions come look in some ways rather like the rules of a generative grammar. The relation between forme de fondation and forme fondée is something like the relation between input and output of a rule, and proportions where terms are 'virtual' in that they abstract away from certain overtly present surface features are comparable to rules which apply to abstract representations at nonterminal stages in derivations. For example, (3a) and 3b) can be construed as representing the generalization of ablaut rules at a point in the synchronic derivation of Indo-Iranian and Germanic which precedes the application of vowel contraction and of the rules which determine the distribution of syllabic and nonsyllabic sonorants. Indeed, Kuryłowicz (1964) has explicitly justified the abstract character of analogical change by the principle that redundant features or elements are to be disregarded in setting up proportions. In our terms that translates into the claim that rules can generalize even when followed by other rules which add those redundant features or elements. Thus, Kuryłowicz (followed by Watkins (1969)) considers the proportion (2) a valid account of the replacement of Indo-Iranian mid.3. pl. *ra* by *nta*:

(4) act. 3.sg. *t* : 3.pl. *nt* = mid. 3 sg. *a* : X (X = *nta*)

Here the suffix *-t* of the unmarked finite verb form (act. 3.sg.) is to be stripped away, as it were, in order for the terms to match properly.

Still, even in their modified forms, both theories retain many inadequacies. Either side can easily adduce examples of indubitable analogical changes which do not agree with the claims of the other. In fact, much of the debate around analogy over the last ten years or so has consisted of generative linguists showing up various problems for the proportional theory (such as analogical changes that depend in various ways on abstract properties of grammatical structure--significantly more abstract than even the work of Kuryłowicz assumed) and the opposition countering with cases of analogical processes which complicate grammars rather than simplifying them. A detailed assessment of some of this

evidence is presented in the fuller version of the present work. It fully confirms Hock's (1975) conclusion that analogy can involve different levels of abstraction: there are some changes 'which can be plausibly explained or motivated only in terms of (the generalization of) sometimes fairly abstract rules, and not (or only with great conceptual difficulties) in terms of classical surface-oriented analogy', and other changes which show a sovereign disregard of those abstract rules and are 'explainable only in terms of surface analogy'.

Because of space limitations I shall omit entirely any discussion of the first type of case, which contradicts recent proposals to revive surface proportions and reaffirms the fundamentally grammar-based, simplificatory character of analogical change. What follows is a brief outline of the problems posed by cases of the second type, cases which appear to show just the opposite. I shall argue that they too can be analyzed as simplifications provided we take into account certain kinds of interaction between the language learner and his linguistic environment and certain kinds of inertia inherent in the process of language acquisition. That is, the difficulties turn out to be artifacts due to the simplifying idealizations of the ideal speaker-hearer and instantaneous language acquisition, which are appropriate in the study of formal grammar but should not be imported into historical linguistics.

Cases where analogy produces complication in the grammar can be grouped into two main types, not necessarily mutually exclusive or in all cases clearly separable, but still broadly distinct in their symptoms and causes: partial analogy and false analogy.

2. *Partial analogy.* The characteristic feature of partial analogy is that an analogical process which would be a simplification if it took place across the board ends up actually complicating the grammar because it is not fully carried through. Each type of simplification can be partial in this sense. As a counterpart to the loss of a rule (and in fact usually as the first step in that process) we find its curtailment, i.e. partial loss. For example, Sadock (1973) finds that the loss of the word-final devoicing rule in Yiddish went via an intermediate stage at which it applied in a phonologically restricted context. 'First, it was lost everywhere except in the final cluster- *-nd*, and then after the loss of final schwas, the rule was dropped entirely.' Thus 'the Yiddish word-final devoicing rule in all likelihood was made considerably more complex (i.e. it became more restricted in its application) before it was lost' (p. 793). Other cases have been presented in Andersen (1969), Dressler (1972), Ralph (1975:161 ff). As a counterpart to the generalization of a rule, we find its partial generalization in some subenvironment only (Thomason 1976; Ralph 1975:172ff). As a counterpart to reordering, we find a new order of application introduced for only some of the cases (Kiparsky 1973:93-101; Robinson 1976; Bley-Vroman 1975).

In such cases, the old form of the rule, or (in case of reordering) the old order remains in the grammar and continues to be applicable to a subset of forms, which must be identified by a contextual restriction added to the rule, or by an arbitrary rule feature marked on some lexical entries--and at worst, by a combination of both. The regularization may,

for example, go through only:

--in inflection as opposed to derivation
--in some morphological category (e.g. plurals but not singulars)
--in some inflectional subclass (e.g. the most productive declension or conjugation)
--in some subset of the vocabulary (often the less common words)
--in some phonologically definable circumstances (often in cases which are in various ways relatively less salient)
--in just one sense of a word (often e.g. in metaphorical use)

In all such cases, the result is a more complex grammar.

Can even such changes be due somehow to the mechanism of simplification? The obvious possibility to consider is to reassess the form of grammars or the evaluation measure (or both) in the hope of discovering that the grammars which are actually learned by speakers are such that even in the apparently 'bad' cases the new grammar turns out to be more highly valued after all. Methodological considerations would recommend this approach because it would give both the simplest and the most restrictive theory of analogy--provided of course that the required kinds of synchronic grammars could be justified in their own right independently of the historical changes which we wish to explain.

Several such attempts have in fact been made. One is to maintain the evaluation measure unchanged while assuming grammars of a much 'shallower' type than generally envisaged. The best-known solution of this type is embodied in the theory of Natural Generative Grammar (Vennemann 1974; Hooper 1976; cf. the related ideas of Skousen 1975). As Hooper notes, the gradual spread of certain kinds of leveling (rule loss) through the lexicon could be characterized as simplification if phonological rules and representations were made very 'concrete'. If, for example, a morpheme was entered in the lexion as the set of its allomorphs, then any analogical change which reduces the number of allomorphs--all cases of rule loss, for example, are of this type--would entail a simplification in the phonological representation of any item subject to the analogical change. If marked ordering of phonological rules did not exist, then for example what now are treated as counterfeeding orderings would really be lexical exceptions to the counterfed rule, and a 'reordering' into feeding order would really be the elimination of this exception feature, which one would expect to take place word by word, as does any idiosyncratic lexical marking. If we could assume that subrules are not collapsed as is claimed in the theory of generative grammar, but are to be kept separate, then obviously the independent generalization of those subrules would not be a problem. In short, the fragmentation observed in the problematic cases of change could be ascribed to the synchronic system in which the change takes place.

This line of thinking, however, destroys the basis for understanding the normal situation where a rule is generalized in the same way in different subenvironments. The splits effected by partial analogy are so varied that in order to turn them all into simplifications we end up with no synchronic generality at all, so that each individual case of alternation is governed by a 'rule' of its own. The result is a synchronically wrong description and a treatment of analogy which suffers from the same kinds of defects as the proportional model.

A more promising-looking attempt to motivate partial analogy in synchronic theory is sketched out in Andersen (1969). Though he too proposes fairly shallow underlying forms, he explicitly assumes single (nondisjunctive) representations related to the phonetics by extrinsically ordered rules, among which he distinguishes morphophonemic and phonological rules. The key to his solution is not so much the 'concreteness' of grammars as the idea that at the stage of the language which immediately precedes the appearance of the analogical forms, speakers learn as it were the 'wrong' grammar, which is such that the subsequent spread of the analogy in fact constitutes a simplification. The scenario for word-by-word extension of a rule would then be as follows: at some point language learners acquire the rule in overgeneralized form. 'When this is the case, lexical items which for historical reasons are not subject to the rule (i.e. do not constitute an environment for its application) have to be specified as exceptions to the rule. The extension of the domain of the rule consists in gradual elimination of such exceptions. '. . .'The first phase can be understood as motivated by the need to formulate grammatical rules in as simple (general) terms as possible. The second phase follows as the traditional forms are imperfectly passed on to succeeding generations of learners. It entails a simplification of the lexical component of the grammar. The first phase is necessarily abrupt--but it is covert and goes unnoticed by the speakers. The second phase is gradual and may be more or less apparent to the speakers as older and newer forms are utilized as indices of style' (p. 822).

In order to account for word-by-word curtailment, Andersen then supposes that speakers can also undergeneralize: 'If we suppose that this rule in its original formulation applied only to morphemes specifically marked as subject to the rule, then we can understand its curtailment--and eventual elimination--as the result of a simplification of the lexicon, exactly analogous to the one we defined for the extension changes. Here then, as in the extension changes, we have two phases: first a morphophonemic rule is formulated in such a way that individual lexical items have to be marked with respect to that rule; then the marking of these items is gradually omitted. The difference is that where the marked items are exempt from the rule, the simplification of the lexicon results in an extension of the domain of the rule; but where the marked items are the ones that are subject to the rule, the result of a simplification is a curtailment of the domain of the rule' (p. 823). The idea is, simply put, that language learners failed to discover the context of the rule and consequently had to mark the forms subject to it with an ad hoc rule feature, whose elimination case by case curtails and eventually liquidates the rule.

The interesting feature of this theory is that it claims to predict the direction in which a rule will change. Whether a rule is extended (eliminating the exceptions to the rule) or curtailed (eliminating the rule) is determined by its status (as major or minor, motivated or unmotivated, etc.). As things stand, this apparent explanatory power is dubious because there is no theory that tells us independently of the change itself what the relevant character of the rule in fact is. It is well enough recognized that the productivity of a rule is not at any rate a simple function of the frequency with which it is applicable (whether

counted by types or tokens). Rules which are originally restricted to a handful of forms may spread, and conversely, dominant rules may lose out. Andersen also appeals to the naturalness of a rule: phonetically motivated rules are extended but phonetically unmotivated rules are curtailed. This is beset with the same circularity because the status of a rule as motivated or unmotivated may depend on the choice of distinctive features, which in turn is made by looking at the subsequent changes that depend on it.

It would appear that this solution does not remove the indeterminacy but simply shifts it: instead of determinate synchronic systems but unpredictable alternative possibilities of eliminating points of linguistic complexity, we have indeterminate synchronic systems of which each alternative sets a fixed path for future analogical change. In effect, this is an attempt to reduce partial analogy to false analogy.

There are difficulties with extending Andersen's theory to partial analogy in general. It does not work where the subset of cases affected is structurally (phonologically or morphologically) rather than lexically defined. At the intermediate stage of the loss of final devoicing in Yiddish, if Sadock is right, the domain of the devoicing rule was cut back from all obstruents to just *-nd*. The rule at this stage has added on a phonological context which unarguably makes it more complex than before. Even if we suppose that the rule had, before its curtailment, somehow become a minor rule (which in itself does not seem a justifiable assumption), it surely cannot be maintained that adding contexual restrictions even to minor rules makes them simpler! Only the kind of curtailment that comes from eliminating features on lexical items that mark them as subject to minor rules can be plausibly regarded as simplification. It is also impossible to suppose that the devoicing rule itself remained in its original general form and the curtailment was effected by marking lexical items as exceptions to it. Apart from the difficulty that the prior status of devoicing as a minor rule again cannot be justified, this would amount to saying that it is only accidentally true of the words which maintain devoicing at the intermediate stage that they all end in *-nd*. How could such a regularity have arisen if it was not apprehended as a rule by language learners at some point? The possibility of a purely item-by-item specification is excluded still more clearly in Thomason's Slavic example. The class of lexical items to which the rule applies at the two stages must be specified in a general way in the rule itself since the class of Animate nouns is an open one. It then follows that the generalization of the Accusative = Genitive rule results in a more complex rule.

A second general type of approach is keeping more or less orthodox assumptions about the form of grammars and instead looking for some other version of the evaluation measure which would reduce instances of partial analogy to legitimate optimalization.

A plausible scheme for revising the evaluation measure might be to keep simplicity in it but to have it interact with other parameters in some way. One such proposal, due to Ralph (1975:179), is to adopt Chen's (1973, 1974) device of 'metarules' as a regulator of partial analogy. These are akin to the 'marking conventions' introduced by Chomsky and Halle (1968:Ch. 9) to account for the 'intrinsic content' of the features, and

still more to the scales of environments utilized by Foley (1977 and many earlier writings), but they implement the idea in a rather different way. Suppose we assign a higher 'index' to the fronting of long and rounded vowels than to the fronting of the corresponding short and unrounded (or lower) vowels, e.g. for the special case that is relevant to Old Swedish palatal umlaut as discussed by Ralph:

(5)

2	ɔ	ɔ̄
1	a	ā
	1	2

Then multiplying (or adding) the indices for each feature will give a ranking of multifeature contexts which corresponds to their historical order of appearance:

(6)			
	a	1	palatalized first
	ā,ɔ	2	palatalized next
	ɔ̄	4	palatalized last

Given a general theory of 'metarules', the order of generalization would be predicted by the principle that rules are generalized in the order of increasing indices.

In one respect this approach represents a major advance over the others we have considered so far: it implies, correctly, that the pattern of partial generalization is not random. It makes the interesting and, I suspect, correct prediction that no language will be found in which such a fronting process has unfolded in the reverse order, although from the viewpoint of the classical simplicity measure that order would be exactly equivalent to what actually happened in Old Swedish.

No doubt the metarules as conceived by Chen and adopted by Ralph are lacking in generality. To assume that a table like (5) needs to be specially set up to index the possible contexts of every phonological process is to assume that there are no general principles behind the ordering of contexts, which is scarcely credible. Many cases fall under the rather simple overarching principle that assimilation proceeds in the order of increasing phonetic 'distance' between focus and environment, as measured by the number of relevantly shared feature specifications between them. The more alike two segments are, the more prone they are to assimilate, other things being equal. For example, velars will be assimilated (palatalized) by an adjacent [-back] segment in the following order: palatalized consonants, *y*, *i*, *e*, æ. By the same principle, *i* fronts *a* before it fronts ɔ. I do not wish to claim that this is the only principle at work: I would suppose, for example, that there are other reasons why the short vowels are fronted before the long vowels. Perhaps the explanation for the ordering of contexts will turn out to be exceedingly complex. This would still be better than a list of 'metarules', which can at best claim to be a summary of the observed facts but not an explanation.

Moreover, neither metarules nor even a more general set of principles which would predict the order of contexts would yet address the entire problem of partial analogy. It is hard to see how these ideas could be

extended to cases of partial analogy in morphological subcontexts, and the lexical cases are evidently quite impossible to handle that way.

For the morphological and lexical cases there does exist another proposal which is entirely different from metarules in content but resembles it formally in that it involves enriching the evaluation measure. This is the idea of Harris (1973), Kiparsky (1972) and others that paradigmatic uniformity has to be recognized as an independent factor in change, not instead of but alongside simplicity (cf. Schindler 1974 for discussion). In effect, it says that the evaluation measure is to somehow check the outputs of grammars (it is impossible to tell just from the rules) and assign a 'cost' to allomorphic variation in paradigms. That is, it claims that the language learner preferentially selects grammars whose outputs have certain 'transderivational' properties. This is a retreat to a much weaker theory, since evaluation measure otherwise is defined purely on the grammar. It is also distressingly vague because crucial aspects of the proposal were left in the air: is allomorphy to be quantified by types or tokens? What are paradigms? How are the relative contributions of allomorphy and complexity in terms of the number of feature specifications (not to speak of opacity) to be weighted relative to each other (King 1972, Kiparsky 1974)? In itself, that does not mean that the idea of a complex evaluation measure relying on allomorphy or some equivalent concept as one of its several components is necessarily wrong. The vagueness could be eliminated by exact definitions of 'allomorphy', 'paradigm', and other key notions of the theory, and some of it may indeed have to be retained as an inescapable correlate of the indeterminacy of analogical change itself. A priori there is no reason why there should not be independent criteria of evaluation in as functionally complex a system as language is, and why their relative weight should not be in certain respects indeterminate--all the more so if it can be shown that the weighting varies with linguistic function, so as to create styles adapted to the special requirements of speaker, hearer, and learner (Kiparsky 1975).

There is another general consideration which ought to give one pause. The status of 'paradigm uniformity' and other such criteria in the theory of generative grammar would be quite different from the role of the corresponding notions of analogy in the neogrammarian theory. In the latter, proportional analogy is assumed to underlie language acquisition and speech production, and its rule as a mechanism of change is simply a reflection of its more basic function. In the theory of generative grammar, simplicity has a corresponding role in language acquisition, and its relevance to analogical change would therefore be expected. Proportions, paradigmatic uniformity etc. do not figure in the theory of generative grammar and it would therefore be surprising to find them directing change. It would mean that there would be no essential relation between the way language changes and anything else about it.

Although some authors (e.g. Lass 1976) have welcomed the proposal to introduce paradigmatic notions into the evaluation measure, I believe Bever and Langendoen (1971) were closer to the mark in criticizing it, or rather the program of which it formed a part, for treating function as structure: that is, for attempting to explain within the linguistic system a phenomenon which originates in the interaction of that system with other systems. The

commonsense explanation for the uneven progress of analogy is that speakers resist some kinds of innovations and readily adopt others. To make something of this idea we surely have to go beyond the formal theory of grammar to the theory of language use, specifically to that part of it which concerns the way speakers deploy the variation deposited in their grammars by ongoing linguistic change (Labov 1972). I shall consider here one way in which such a program might be worked out.

One could picture partial analogy arising through the interplay of language acquisition and synchronic variation by a process somewhat analogous to that by which the conditioning of sound changes is crystallized out of their initially variable application. A three-part scenario for it would be:

(1) Language learners can acquire different grammars even when exposed to the same language, because of random differences in the speech data they are exposed to, in the order in which they are exposed to it, in the way they analyze indeterminate cases, and possibly because of other reasons. Such 'imperfect learning' (along with language contact and other sources of linguistic innovation) constantly replenishes the language with a fresh supply of coexisting variants.

(2) Some of the variants arising this way will tend to be avoided in speech because they are dysfunctional or because they become stigmatized. Others will be favored in speech because they can be used to make things easier to say or to understand. These effects will show up in varying degrees depending on the circumstances.

(3) Some of the bias in the deployment of variants will in turn be grammaticalized, with favored variants being acquired as obligatory and disfavored variants being lost.

The assumption is, in short, that the variation due to 'imperfect learning' which we postulate as the source of analogical change is channeled jointly by the conservative influence of the established norm and by the functional needs of the system. Their combined effect should, if the assumption is right, predict whatever is predictable in the patterns of partial analogy. Let us consider the two factors in turn.

The first way an analogical innovation might end up incompletely effected would be, then, that it gives way in particular subenvironments and/or in particular lexical items to the older standard, whether through self-monitoring by language learners, through explicit correction by other speakers, or both. Would any general type of pattern be likely to emerge from such a process? Presumably there could be no completely predictable pattern, if only because the linguistic data which bring about the partial adaption of the older system will come to the attention of different individuals in different orders and at different stages in their evolving grammars. Still, one would expect some innovations to stick out against the norm more than others, and thereby to have a diminished life expectancy. The innovations most likely to succeed would be the least 'blatant' or 'salient' ones, for they have the best chance of surviving the language learners' own efforts to accommodate their speech to the community norm,

and are also least likely to provoke outright correction or ridicule by other speakers.

Intuitions on what exactly defines the crucial dimension of relative saliency are bound to be vague. Fortunately we have some slight empirical foothold on it, for dialectologists have long operated with the notion of relative saliency in studies of interdialectal adaptation.

The relevance of certain kinds of saliency in syntactic change has recently been demonstrated in an important article by Naro and Lemle (1977). They investigate the loss of subject-verb agreement in Brazilian Portuguese and conclude that the process spreads through the language in order of increasing saliency. They take the relative saliency of an innovation against the background of the old form to be a composite function of distinct factors, of which they name two:

(1) An innovation which is more differentiated from the old form is more salient than one which is more similar to it.

(2) An innovation is more salient in monitored (formal) speech than in unmonitored speech.

It is plausible to add to this the following factors:

(3) An innovation in a frequent form is more salient than an innovation in a rare form.

(4) An innovation which alters surface structure or phonotactics is more salient than one which does not.

This permits us to interpret many types of partial analogy: the tendency for 'small' alternations to be eliminated before 'big' ones (1), stylistic differentiation (2), frequency (3), and various 'structure-preserving' effects (4).

In addition to an overall normative pressure which pushes back particularly the most salient innovations, we are also supposing as another selective process, in line with traditional functionalist thinking, that innovations which serve the needs of speakers well or ill will be resisted or favored accordingly. Here is where the familiar functional factors find a natural interpretation.

Obviously the factors that control the use of variants in speech, whatever they turn out to be, can have nothing to do with analogy per se, for variation from <u>any</u> source must be patterned in the same way once it is in the language (Samuels 1972:87). This is borne out by the resemblances in the characteristic profiles of partial analogy and variable rules (Kiparsky 1972). We have in effect ended up dividing the explanation of analogical change between two different theories, one dealing with the <u>source</u> of the innovations (imperfect learning) and the other with their <u>selection</u> in speech and eventually in grammar. (A vague parallel might be the theory of evolution as a function of genetic variation and natural selection.) The constraints on analogical change are, then, jointly

determined by the properties of imperfect learning and of variation.

This part of my proposal is close in general spirit to Samuels (1972), who also distinguishes between the components of 'variation' and 'systemic regulation' in change (see especially p. 135 ff). He remarks that 'neither sound-change nor analogy can be regarded as providing any more than the raw material of change; the process of change, from its initiation to its acceptance in a system, is considerably more complex than this . . .' (p. 136). Samuels devotes his attention mainly to 'systemic regulation' and rather underestimates the structural conditions of change. I cannot agree with him, though, that new variants always reduce information-content and that subsequent selection is a 'compensatory mechanism' which 'consists of the restoration of the distances lost in [variation]' (p. 178).

A certain amount of discussion has centered on the question whether change originates in competence or performance (see Dressler 1976 for recent discussion). From the present point of view, the question is rather misleadingly posed, since change involves both competence and performance. This is also the conclusion reached through a different line of reasoning by Vincent (1974:436): '. . . analogy is neither a competence nor a performance phenomenon, but rather a sort of line linking these two aspects of language.'

3. *False analogy.* There is a second fallacy in assuming that the retention of forms from intermediate stages of language acquisition must simplify the grammar: they might be residual projections from an analysis which, even though optimal for the restricted linguistic data avaliable to the learner at some stage, is not part of the optimal grammar of the whole language. If they are retained anyway, the result is analogy which creates a complication of the grammar. Everyone knows how the crucial data that may decide between competing analyses can lurk in obscure corners and elude the linguist working from a limited corpus. It would not be surprising to find the language learner occasionally in the same predicament. Accordingly, the assumption that analogy must represent simplification with respect to the adult grammar is entirely unwarranted. The most that we can legitimately claim is that analogical innovations represent analyses which are optimal at the particular stage of language acquisition at which these analyses arise and for the particular set of primary linguistic data which is under consideration by the language learner at that state. Even a partial retention of these innovations may force a complication on the grammar when it is reanalyzed on the basis of a fuller set of data.

It is obvious that the sort of paradigmatic leveling which fails to simplify and possibly complicates the adult grammar is often ascribable to a simplification at an earlier stage in language development. For a learner of Latin whose relevant data includes only the oblique case forms *honōrem*, *honōris* etc., *honor* is the simplest projection for the nominative singular. Never mind that having later learned *honestus* and acquired an intervocalic *s* → *r* rule anyway for cases like *flōs∿flōris* he might be better off with the older *honōs* after all: the horses are already out of the barn. For a learner of Canadian English who has not acquired the *ay* ∿ *əy* and vowel shift rules, *write* would simply be underlying /rəyt/, and the simplest projection for the suffixed forms *write* + *er*,*write* + *ing*

would be [rəydər], [rəydIŋ]--though once the dependence of *ay* vs. *əy* on the voicing of the following consonant is learned, the transparent ordering which yields the older forms [raydər], [raydIŋ] might be preferable.

Similarly, cyclic ordering can be a structural accommodation of analogy which imposes on derived words the surface features of stems they contain. It is not necessary, if this is so, to suppose that cyclic ordering is to be assigned preferred status by the evaluation measure, for the tendency of rules to shift into the cycle can be explained without it.

This inertia of the language learner's own mistakes is perhaps the minimal assumption which can be made to account for 'false' analogy. It should be distinguished from McCawley's (1975) more radical, by no means implausible idea (cf. Fodor 1978:62) of an inertial effect in the language learner's own grammars. One version would be that language learners resort to restructuring only 'under duress', perhaps if the grammars they have formulated cannot be patched up to incorporate new data in a way which satisfies some criteria--which remain to be specified by a theory of language acquisition. Alternatively, one could imagine the degree of resistance to restructuring depending how great the restructuring would be, perhaps as weighed against the amount of simplification it would afford. These hypotheses would entail that the evolving grammar is not necessarily at any particular stage, including the finished grammar, optimal for the data it covers, though it of course might be. Even more severe would be the assumption that all language acquisition is, in effect, a patchup job, with new data always being absorbed by means of the smallest possible grammatical adjustments, without restructuring (cf. also Haber 1975 for some related observations). That would mean that the child's first hypotheses always fix the future course for the evolving grammar. The optimal grammar would on this view be an ideal construct which was perhaps never realized by any actual language learner.

Similar in spirit is Andersen's (1973) idea of 'adaptive rules'. It amounts to assuming that the learner may be stuck with a nonoptimal analysis which in itself generates the wrong output and is mapped into the right one by 'adaptive' rules. Andersen illustrates his concept with an interpretation of a sound change which takes palatalized labials to dentals. He assumes that at some point language learners began to interpret palatalized labials as being dentals ([p'] as /t/ etc.). To account for their actual pronunciation as [p'] rather than [t] they therefore required an adaptive rule which turned dentals to palatalized labials (/t/ → [p'] etc.). Having thus become 'virtual' dentals by reanalysis, the palatalized labials changed to dentals phonetically by the loss of the adaptive rule. I admit to finding this unconvincing as a proposal about sound change. The thought that language learners would somehow come to analyze palatalized labials as being really dentals, without any phonological motivation for doing so, and later start to pronounce them in accordance with this reanalysis is too idealist for my taste. For analogical change Andersen's notion of adaptive rules might make more sense. Even there, however, they seem to me to underestimate the flexibility of the language learner.

These hypotheses make different but largely overlapping predictions about analogy. They share the essential property of allowing for the

real-time course of language acquisition to determine aspects of the system that eventually emerges. The precise causes of the inertia, and the degree to which it is mediated by the structure itself, could well be some combination of those mentioned here. The essential empirical difference between my assumption that the language learner freely restructures his grammar when encountering new data that motivates it, and McCawley's and Andersen's assumption of more limited adaptive measures for coping with new data, would seem to be the following. On the assumption of continuous restructuring, a certain amount of selection and organization would be expected to be imposed on the innovations by the adaptive system, due to the efforts of the learner to integrate the new data into his grammar in the optimal way. If on the other hand the new data is accommodated by tacking on 'patchup' rules or 'adaptive' rules, to the already formulated grammar, no such structural effects should presumably show up. Herein lies the relevance of cases where the result of analogical leveling is systematically adopted into the language precisely where it can be accommodated in a simple way into the system, and suppressed elsewhere. This would be incomprehensible if we did not assume grammatical reintegration rather than makeshift 'adaptive' rules.

'Adaptive' rules do appear to be the right mechanism for certain sorts of cases, hypercorrect forms, for example. They show precisely the distinction between the ordinary result of language acquisition, where reanalysis takes place, and the special situations caused by limited access to the relevant linguistic data, which lead to hypercorrection.

If this view of paradigmatic leveling is right, then it follows that efforts to combine paradigmatic leveling and 'distinctness' effects in terms of a unified set of 'paradigm conditions' (Kiparsky 1972) or 'Humboldt's Universal' (Vennemann 1974) were fundamentally misguided. Paradigmatic leveling is a direct reflex of language learners' tendency to simplify grammar, whereas avoidance of homonymy, distinctness of grammatical categories etc. are imposed on the system by the purposive selection of speech variants from whatever source, including ordinary sound change. As a result of their different etiology, they are also implemented in the grammar in different ways. It is striking that 'distinctness' may be effected by the most varied structural means. Even for one category within one grammar it is possible to demonstrate the characteristic 'conspiracy' situation where the distinction between morphological markers is maintained by numerous rules and restrictions on rules acting in concert (cf. Kiparsky 1972:214). Hogg (1975) has pointed out that such cases do not seem to occur with paradigmatic leveling. This is predicted from our assumption that the selection of variants favors category distinctness but is neutral as to paradigmatic leveling.

We have argued that 'paradigm' conditions can be eliminated from the evaluation measure. It is tempting to try a similar reduction for opacity. Can we interpret shift into transparent rule order as the result of incorporating into the language the projections of the simplest grammar for some set of core data, as we did for paradigmatic leveling? That would let us pare down the evaluation measure even further by taking even opacity out of it, essentially restoring it to its original form of a pure simplicity measure.

Let R_1, R_2 be two opaquely ordered rules which are reordered into the transparent order R_2, R_1. Consider now how the reordering can be analyzed as the grammatical reintegration of the projection from R_1 at a stage when R_2 has not been acquired. Taking first reordering into feeding order, suppose we have the rules

(7) i. a → b / _ c
ii. d → c / _ e

where (ii) counterfeeds (i), e.g. /ade/ → *ace*, not **bce*. Suppose that the primary linguistic data under consideration by the language learner at stage S_n is enough to motivate (i) but not (ii). Such instances of the output *ce* from (ii) as are encountered are therefore incorrectly analyzed as derived from /ce/ rather than /de/ at S_n. From *a*→*ce* the grammar therefore projects the 'wrong' output *bce* by (i). Suppose that at stage S_n+1, rule (ii) has been learned. The wrong output *bce* of S_n can be reintegrated into S_n+1 simply by ordering the rules in the transparent order (ii,i). If the analogical *bce* wins out, reordering has taken place. If correction to *ace* prevails, normal language transmission has taken its course. So it is possible to account for reordering into feeding order by the process of grammar simplification (imperfect learning) <u>without</u> supposing that feeding order, or transparency, is in itself more highly valued, if the entirely natural additional assumption of learners' output inertia is made and we further postulate an order of acquisition. Note that our scenario <u>predicts</u> the asymmetry of reordering. A switch from feeding to nonfeeding order would be inexplicable by the process we have outlined.

The above proposal is extreme in that it reduces assumptions about the factors governing language learners' structuring of data to the SPE minimum (simplicity) and puts the whole burden of explaining transparency effects, as was done with leveling, on the postulated inertia of learner's own earlier speech forms and on assumptions about the order of acquisition of rules. There is also a more moderate version of the same general idea, where the role we have so far assigned to the order of acquisition of rules is given to a minimal additional assumption about the language learners' structuring of data:

(8) For any set of data D, learners select the shallowest underlying representations compatible with the simplest grammar of D.

All that this says is that speakers do not indulge in gratuitous abstraction over and above what is required for linguistic generalizations. (8) is of course quite distinct from such constraints as the Alternation Condition, which limit the abstractness of underlying representations <u>regardless</u> of any resulting grammatical complications. It has, I think, been generally accepted even by those who reject these severer constraints, and would seem to be independently necessary even if these are accepted, in order to predict the restructuring on nonalternating outputs of neutralization rules (for example, word-final obstruents restructured as unvoiced in isolated forms in languages like German, Yiddish, etc., with final devoicing rules, where the Alternation Condition does not predict the choice of the devoiced obstruents as basic because the devoicing rule applies in a derived environment).

Most of the predictions about transparent ordering can be had this way too. Instead of assuming a stage where only (i) has been learned, suppose that (ii) has been learned but that those outputs of it (*ce*) for which the learner is not (yet) in a position to motivate an underlying source /de/ are systematically represented by him as /ce/, even though the more abstract /de/ would be equally simple. (So far we have done without this assumption). Then the grammar projects in such cases from *a+ce* the output *bae* rather than *ace*. If retained, it must be structured in terms of feeding order in all cases where *ce* is reanalyzed by the learner as /de/ on the basis of data encountered later.

Another likely victim of our reductionist rampage is the revised Alternation Condition (Kiparsky 1973), which stipulates that obligatory neutralization rules (or, in a milder formulation, obligatory nonautomatic neutralization rules) apply only to derived inputs.

4. *Conclusions.* To summarize, we have discussed two recalcitrant types of analogy, partial and false analogy, and concluded that they too yield to the interpretation of analogical change as simplification under a more realistic interpretation of change which allows for inertia in the acquisition process and for interference between the learner's system and the target norm of the speech community. This is partly good news and partly bad news. The bad news is that linguistic change can no longer be used as a probe into linguistic structure in the direct and naive sense that many of us used to think. Before we can exploit historical evidence for synchronic purposes we need a firm theory of the intervening factors--the effect which the concrete process of acquisition itself and differential saliency can have on what is acquired. It follows for example that the arguments concerning abstractness will have to be rethought in so far as they are based on historical evidence. The good news is that we can restore the integrity of the evaluation measure by eliminating from it the accretions which it had acquired. These accretions have two suspicious properties in common. First, they require reference not to the form of the grammar (as simplicity does) but to the relationship of the output of the grammar to something else--either to other output forms (paradigm conditions), to rules (opacity), or to underlying forms (recoverability, Hale's deep/surface disparity avoidance). Second, they were motivated practically exclusively from historical evidence. Arguments that they are required to select the synchronically correct grammar have not been forthcoming. The elimination of these considerations from the evaluation measure would therefore be a gain rather than a loss.

We wind up with a surprising conclusion. There are two assumptions of the theory of generative grammar that have been subjected to strong criticism: the simplicity metric, and the idealizations of instantaneous language acquisition and the ideal speaker-hearer. What we have seen is that by eliminating the second from historical linguistics we overcome the objections to the first. We seem to have found a way of retaining the simplicity measure without giving up what is right about traditional functionalist thinking. Structural and functional considerations both play an essential role but as part of separate systems. This relieves the evaluation measure of the burdens which purely structural theories unjustly placed on it and eliminates the unfortunate 'teleological'

element of traditional functionalism, as well as its inability to account for structural patterning in analogical change. The 'teleological' effect can now be seen to follow from the functionally and structurally determined 'filtering' of the variation generated by the process of imperfect learning, while the structural patterning follows from the source mechanism of imperfect learning as well as from the grammaticalization of the functionally patterned variation.

NOTES

[1]This essay is extracted from a monograph on analogy to be published separately. It presents the main line of argument in a condensed form, omitting most of the critical discussion of alternative views and nearly all supporting data. This work was supported in part by the National Institutes of Mental Health Grant No. 5 P01 MH13390-12.

[2]There is a terminological confusion to be cleared up. The term analogy has become established in reference to both a particular type of phenomenon in linguistic change, whose existence is not in dispute, though there may be disagreement about where its exact boundaries lie, and a particular theoretical interpretation of this phenomenon, namely the proportional schema or its equivalent, whose correctness has certainly been cast into doubt. My own practice earlier was to simply avoid the term, and to replace it, in both senses mentioned above (the phenomenon and the process I was proposing to account for it) by simplification. The term is also avoided by Andersen (who likewise objects the traditional interpretation and speaks of simplification (1969:824), though from a somewhat different point of view): 'The changes with which we will be concerned would be called analogical by some. We will not use this term, for terms like "grammatical analogy" as traditionally used are highly ambiguous, and we need to be specific' (p. 807). Others have continued to use the term analogy to refer to the mechanism as traditionally understood, that is, to the extension of surface patterns by the proportional schema or its equivalent (e.g. Chomsky and Halle 1968:156, fn. 12; Isačenko 1970). Finally, King (1969) and others maintain the term in ambiguous use both 'as a cover designation for . . . instances of change' (p. 128) and as a term for the proportional mechanism (e.g. p. 182). The ambiguity would seem to be easily resolved by context but has nevertheless spawned a fatuous line of argument where generative criticisms of the surface proportional theory of analogy are called 'attacks on analogy', and their authors are then, absurdly, denounced for failing to distinguish between analogy and sound change!

None of these ways is satisfactory. It is best to have a separate, neutral term, for the type of historical change whose interpretation is at issue, and distinct terms for the various processes which have been proposed to account for it. I will here use analogical change (or sometimes just analogy) for the historical phenomenon and proportional analogy for the general class of processes by which it is traditionally explained (suitably qualified when necessary to devote specific versions of it), and simplification for the general class of processes which has

been more recently suggested in its stead (again suitably qualified when necessary).

REFERENCES

Andersen, H. 1969. A study in diachronic morphophonemics: The 'Ukrainian prefixes'. Language 45. 807-30.

------. 1973. Abductive and deductive change. Language 49. 765-93.

Bever, T., and D. T. Langendoen. 1971. A dynamic model of the evolution of language. Linguistic Inquiry 2. 433-64.

Bley-Vroman, R. 1975. Opacity and interrupted rule schemata. Papers from the Eleventh Regional Meeting, 73-80. Chicago: Chicago Linguistic Society.

Chen, Matthew. 1973. On the formal expression of natural rules in phonology. Journal of Linguistics 9. 223-49.

------. 1974. Natural phonology from the diachronic viewpoint. Papers from the Parasession on Natural Phonology, 43-80. Chicago: Chicago Linguistic Society.

Chomsky, N., and M. Halle. 1968. The sound pattern of English. New York: Harper and Row.

Dressler, W. 1972. On the phonology of language death. Papers from the Eighth Regional Meeting. Chicago: Chicago Linguistic Society.

------. 1976. How much does performance contribute to phonological change? Wiener Linguistische Gazette 13. 3-18.

Fodor, J. D. 1978. What kind of exception is *have got*? Linguistic Inquiry 9. 45-65.

Foley, J. 1977. Systematic morphophonology. Cambridge: Cambridge University Press.

Haber, L. R. 1975. The muzzy theory. Papers from the Eleventh Regional Meeting, 240-56. Chicago: Chicago Linguistic Society.

Harris, J. 1973. On the order of certain phonological rules in Spanish. A festschrift for Morris Halle, ed. by S. Anderson and P. Kiparsky, 59-76. New York: Holt, Rinehart, and Winston.

Hermann, E. 1931. Lautgesetz und Analogie. Abhandlungen der Gesellschaft der Wissenschaften zu Göttingen. Phil.-Hist. Klasse, Neue Folge 23,3. Berlin: Weidmannsche Buchhandlung.

Hock, H. H. 1975. Current trends in historical linguistics. Proceedings of the first meeting of the Kentucky Interdisciplinary Conference on Linguistics.

Hogg, R. M. 1975. The place of analogy. Neophilologus 59. 109-13.

Hooper, J. 1976. An introduction to natural generative phonology. New York: Academic Press.

Isačenko, A. V. 1970. East Slavic morphophonemics and the treatment of the jers in Russian: A revision of Havlik's Law. International Journal of Slavic Linguistics and Poetics 13. 73-124.

Kenstowicz, M., and C. Kisseberth. 1977. Topics in phonological theory. New York: Academic Press.

King, R. D. 1969. Historical linguistics and generative grammar. Englewood Cliffs, N.J.: Prentice-Hall.

------. 1972. A note on opacity and paradigm regularity. Linguistic Inquiry 3. 535-39.

Kiparsky, P. 1971. Historical linguistics. A survey of linguistic science, ed. by W. O. Dingwall, 576-649. College Park, Md.: University of Maryland Press.
------. 1972. Explanation in historical linguistics. Goals of linguistic theory, ed. by S. Peters, 189-227. Englewood Cliffs, N.J.: Prentice-Hall.
------. 1973. Phonological representations. Three dimensions of linguistic theory, ed. by O. Fujimura, 1-136. Tokyo: TEC Company.
------. 1974. On the evaluation measure. Papers from the Parassession on Natural Phonology, 328-37. Chicago: Chicago Linguistic Society.
------. 1975. Comments on the role of phonology in language. The role of speech in language, ed. by J. F. Kavanagh and J. E. Cutting, 271-80. Cambridge, Mass.: MIT Press.
Kuryłowicz, J. 1956. L'apophonie en indo-européen. Wroclaw: Polska Akademia Nauk.
------. 1964. The inflectional categories of Indo-European. Heidelberg: Winter.
------. 1968. Indogermanische Grammatik, Band 2: Akzent-Ablaut. Heidelberg: Winter.
------. 1970. Les verbes germaniques du type *lūkan*. Melanges Fourquet, 157-62.
Labov, W. 1972. Sociolinguistic patterns. Philadelphia: University of Pennsylvania Press.
Lass, R. 1976. English phonology and phonological theory. New York: Cambridge University Press.
Mascaro, J. 1976. Catalan phonology and the phonological cycle. Ph.D. dissertation. MIT.
McCawley, J. 1975. Acquisition models as models of acquisition. Proceedings of the 1974 NWAVE Conference. Washington, D.C.: Georgetown University.
Naro, A., and M. Lemle. 1977. Syntactic diffusion. Ciencia e Cultura 29. 259-68.
Paul, H. 1920. Prinzipien der Sprachgeschichte. Fifth edition. Tübingen: May Ni emeyer.
Ralph, B. 1975. Phonological differentiation: Studies in Nordic language history. University of Gothenburg.
Robinson, O. W. 1976. A 'scattered' rule in Swiss German. Language 52. 148-62.
Sadock, J. M. 1973. Word-final devoicing in the development of Yiddish. Issues in linguistics: Papers in honor of Henry and Renee Kahane, ed. by B. B. Kachru et al., 790-97. Urbana, Ill.: University of Illinois Press.
Samuels, M. L. 1972. Linguistic evolution. Cambridge: Cambridge University Press.
Schindler, J. 1974. Fragen zum paradigmatischen Ausgleich. Die Sprache 20. 1-9.
Skousen, R. 1975. Substantive evidence in phonology: The evidence from Finnish and French. The Hague: Mouton.
Thomason, S. 1976. Analogy as grammar complication. Proceedings of the Second International Conference on Historical Linguistics, ed. by W. Christie. 401-9. Amsterdam: North Holland.
Vennemann, T. 1974. Words and syllables in natural generative grammar. Papers from the Parasession on Natural Phonology, 346-74. Chicago:

Chicago Linguistic Society.
Vincent, N. 1974. Analogy reconsidered. Proceedings of the First International Conference on Historical Linguistics, ed. by J. Anderson and C. Jones, Vol. II, 427-45. Amsterdam: North Holland.
von Wartburg, W. 1969. Problems and methods in linguistics. New York: Barnes and Noble. (Translated from the 2nd French edition of 1963. Originally published in German in 1943.)
Watkins, C. 1969. Indogermanische Grammatik, III.1, ed. by J. Kuryłowicz. Heidelberg: Winter.

MULTILINGUALISM AS OBJECT OF LINGUISTIC DESCRIPTION

Charles A. Ferguson
Stanford University

In this paper I intend to take a reactionary point of view and say something quite old-fashioned. I will say that linguists should write grammars. Then I will add the somewhat radical thought that they should write multilingual grammars. It may take a considerable number of words to say that, but you have fair warning that that is what I am going to say.

Writing grammars is a traditional task of linguists, in fact it is almost a defining task of linguists. It is not the only one, but it is certainly something linguists are known to do. They attempt to provide a principled description of a particular language; either the whole language or some part of it. They have been doing that for a long time, and I am suggesting that they should not lose heart, but should continue doing it, because it is a good thing. To be sure, there are other tasks, such as the formulation of theories about language in general and theories about how to write grammars. Some of these tasks are just as important or possibly more important than writing grammars. But the writing of grammars is a good thing.

Why do linguists write grammars? I would like to think that the first reason they do so is simply to extend our knowledge about human languages, so that we know more things about languages. I realize that in some quarters this is a fairly old-fashioned view of why one would write a grammar, but it is what comes to my mind first. Second, linguists write grammars in order to codify languages. That is to say, to prescribe the form of a standard language for its users. This is also supposed to be not such a good thing according to the ethos of modern linguistics, but if we think back to the great linguists in the history of our science, most of them have written prescriptive grammars. Panini was certainly a prescriptivist, no matter how else we may characterize him, and he was also one of the great linguists and grammar-writers of all time. Third, linguists write grammars in order to have a basis for pedagogy, so that people can teach the language more effectively and once again, although modern linguists belittle such a purpose, some of the great grammarians have had exactly this aim. Sibawihi, the most famous Arab grammarian, certainly wrote his grammar, at least in part, so that Persians would know how to get their Arabic right and so that Arabic could be taught properly to Persians. Next, many linguists are interested in writing grammars so that they can compare one grammar with another, whether for typological, historical, or areal comparison or for pedagogically oriented contrastive analysis. In other words, an important purpose of writing grammars is to learn something about the nature of language in general by comparing grammars. And then finally, I put it in fifth place, you can write a grammar merely to check

out some theory, that is merely to see if you can confirm or disconfirm a particular theory or theoretical model that you are interested in, or perhaps modify the theory or extend it in some way to take care of other kinds of material. People often put this purpose at the head of the list and say that grammar writing is not worthwhile unless you are checking a theory. In the present state of linguistics, I prefer grammars. Good grammars last well and are used by all kinds of theorists, but theories come and go quickly and too often interfere with balanced coverages in grammars.

If we agree that linguists, for these various reasons, do write grammars, whether attempts at whole grammars or partial ones, why is it becoming less fashionable to do so today? I think the chief reason is that the task is getting harder and harder to do. That is, the more that linguists study language, the more they find there is to put in grammars. There is also the basic problem of how to go about writing a grammar, which is becoming increasingly troublesome, too. But the first problem is that there is just so much there. Language is a very complex phenomenon, and linguists are dealing with more and more aspects of it. All you have to think of is how linguistics programs today have full courses in pragmatics, child phonology, social dialects, and a host of topics which a few years ago were not in the catalog. Presumably, these things have to get into grammars, somehow, and it seems to have become an impossible task to write a grammar. In fact, a philosophy of despair has emerged. Some linguists say, 'The most you can do is write some fragment of some possible grammar of some part of some language (preferably with theoretical significance).'

Another problem is the old-fashioned question of what constitutes a language. How do you know when you have one? This is an important part of my topic: natural language as the object of description. It is not easy to define what we mean by a language, i.e., the structural or functional criteria we need for determining when we have a language. Linguists, by and large, have put that problem aside. It comes up every once in a while as something that we have to struggle with, but the fact is we have ready at hand no way of deciding when the object we have is a language, and not a dialect or two languages or something else.

Let us assume that the notion 'a language' has a natural status of some kind as an object of description which is highly productive for understanding the nature of human language behavior in general, how it changes, and how it is processed individually and socially. Such an assumption does not exclude the study of partial, restricted and marginal language behavior which is abnormal or pathological, but it accords with the view that whole languages are the normal object of linguistic description. The criteria for delineating this natural unit would include autonomy, stability, and functional range. By 'autonomy' I mean the degree to which the object in question is structurally and functionally distinct from neighboring phenomena; this may involve such measures as structural coherence, mutual intelligibility, extent of superposed variety, and attitudes of speakers (cf. Ferguson & Gumperz 1960). By 'stability' I mean the degree of internal variation; extremes of diachronic or functional flux in language behavior seem to represent atypical language situations different in kind from normal use of language (cf. Ferguson & DeBose 1977). By 'functional range' I refer

to the degree of restriction in semantic range and acceptable uses; an object restricted to religious use or lingua franca functions and not serving as the mother tongue of a social group is not a full language (Ferguson & DeBose 1977). These criteria are mentioned as illustrations, not as offering a solution. The identification of the linguist's object of description is not a trivial issue; it is a tough problem and one that will not go away just by our ignoring it.

Linguists choose different ways of getting around that problem. One alternative is simply to select a corpus of some kind, and do your best to write a grammar of that. This is usually done not just to describe the corpus itself, but to provide a description that will apply to a larger body of material. Or you can take the way out of just describing one idiolect, and regarding that as the normal object of linguistic description, without attempting to characterize the larger unit. Or you can try to find out what *langue* really is, i.e., find out what is shared by the particular community, and leave out things that are not shared.

There are other ways: you can even back out all together and say 'the language that I am writing a grammar of is whatever it is that my grammar describes', and then the grammarwriter's decisions are inexplicit but not subject to debate. Another thing we can do is to take selected fragments of language and say, 'This is what we are going to write about.' Following this approach, it may be possible to find some kind of useful core fragments that we would all agree are useful information about any language we wanted to describe. Examples of this method are the efforts by Bernard Comrie or the Stanford Phonology Archive to produce a checklist about things you would like to find out about any language that you start to describe (Ferguson 1978:26).

The next issue is variation. We have all come to recognize the fact that languages are full of variation. There is just no way that we can avoid it. An object is not a human language if it does not vary along many dimensions. There must be dialect variation depending upon where the users come from or what part of society they are in. There must be register variation conditioned by the use, the occasion, the addressee, the topic, or whatever. Also, many languages have superposed varieties which we find a little awkward to call dialects or registers, as when a standard variety or even a national standard on top of regional standard is part of a particular language situation. There are actually a number or recognized types of language situations, such as diglossia, standard language with dialects, decreolization continuum, and so on, as well as many situations which do not fit into those types. Most linguists are getting reconciled to the fact that they must include an account of variation in writing the grammar of a language.

Let me give one illustration from my own experience in a diglossic community: the diphthongs in the kind of Arabic that is spoken in Damascus. There are two long mid vowels, $\bar{e}$ and $\bar{o}$, and also two dipthongs *ay* and *aw*, and there is some kind of special relation between them, since many words are pronounced sometimes with the monophthongs and sometimes with the diphthongs. Standard grammars of Classical Arabic give only *ay* and *aw*, and grammars of colloquial Damascene Arabic give $\bar{e}$ and $\bar{o}$ with only marginal

examples of *ay* and *aw*, yet speakers of Arabic in Damascus show extensive variation in the pronunciation of these sounds. Control of the appropriate use of this variation is part of the basic linguistic competence of Damascus speakers and belongs in the grammar. Even a person who has very little education in Classical Arabic has learned from very early childhood to use *ay* and *aw* in certain sterotyped expressions, and even the most scholarly Arab using Classical Arabic pronounces some proper names and foreign loanwords with $\bar{e}$ and $\bar{o}$. What is of greater interest and what forces me to say that the phenomenon cannot be described by describing colloquial Arabic and Classical Arabic and some type of relationship between them, is that it is possible to pronounce most words in the lexicon either way, depending on whether speakers want to show that they are speaking in a classical style or a colloquial style. There is a very complex system of variation in which some words tend to be more classical, some tend to be more colloquial and all fluctuate to a certain extent within the system. An analysis which gives only $\bar{e}$, $\bar{o}$ or only *ay*, *aw* is missing an important point about the way people speak Arabic in Damascus.

There are, of course, various ways of including such material in grammars. I do not have to tell you that people have invented things called variable rules as one way of doing this. There is also an oldfashioned way of describing one part of a language in the body of the grammar and putting the other varieties in footnotes or specially marked passages in the grammar. I am referring to such grammars as Smyth's old Greek grammar (Smyth 1956), which has 'Classical' Attic in the main body of the text and footnotes to tell you the Ionic, the Doric, and so on. Or, let us say, Kellogg's Hindi grammar (Kellogg 1965), which gives an impressive range of dialect variation in the notes and marginal comments.

These may be very primitive attempts to cope with getting variation into the grammar, but sometimes they are much more informative than a grammar which carefully and purposefully removes a great deal of the variation and describes only a uniform, isolated part of the language behavior. One important part of characterizing a language is identifying the nature and extent of variation in it. That is, one does not only compare two languages as to phonological structure, syntactic structure, lexical structure, semantic structure, or other 'structures' or 'components' or 'levels' that traditional linguists talk about. Two languages which are basically similar in phonology and syntax, let us say, but one of which has a great deal more variation of a certain kind than the other, differ in a linguistically significant way. We need ways of talking about the nature and extent of variation in a particular language as a part of its grammar.

Just as variation within a language is universal, we could almost say that multilingualism in speech communities is universal. A great many speech communities are characterized by such structural discontinuities in the kind of language used in the community that by almost any criteria one would want to say that several different languages are used in the community. There are, of course, monolingual speech communities where the variation is so limited that no one would want to speak of more than one language. I suspect we would not have to travel very far from

Champaign, Illinois, for example, to find one of the surrounding towns that is a monolingual speech community in this sense. In such communities there will be register variation, dialect variation, and so on, but it will all constitute unmistakably a single language. However, if we look around the world, including many parts of the United States, there are speech communities in which a number of different languages are part of the linguistic repertoire of the community. Distinct languages exist side by side and are part of the whole scheme of variation of the speech community. What I want to suggest tonight is that if we are going to write a grammar of what goes on in a speech community that uses--let us say--four languages, instead of writing four separate grammars and then writing rules for when people use one language or another, we should try to write a unified grammar in which all this variation fits somewhere. Instead of four separate structures which are variously used, we may have one linguistic structure with complex internal variation.

Perhaps the easiest way to argue for this suggestion is to give examples of multilingual speech communities which linguists have tried to describe. Now, as far as I am aware, linguists have not really tried to write whole grammars of such linguistic repertoires. In fact, what they typically do is to write what is essentially a problem paper. They take some set of interesting facts about a multilingual speech community and say, in effect, 'See, look how these work in this community, and isn't that exciting', or 'interesting', or 'doesn't that have implications for this or that?' But they do not try to say, 'These facts are part of a whole grammar of several languages operating together.' The examples I will use were chosen more or less at random and are not intended to be typical, but they are all possible examples of multilingual speech communities where repertoires could be analyzed as unified grammars.

Norman Denison has written a number of articles about a little town called Sauris or Zahre, an isolated speech community in the Italian Alps (e.g., Denison 1968). It is a little town of fewer than a thousand people, but it has been flourishing for a long time and presumably will continue to flourish for some time to come. Everybody in that town speaks three languages. The local variety of Standard Italian is the language of education, the language of school, even kindergarten, and the language that parents often speak to their children so that they will do better in education. It is the language in which almost all writing takes place, the language of certain kinds of formal uses and so on. Second, there is Friulian, a kind of local regional standard of Italian, which is very different from Standard Italian. This is the language everybody uses informally when non-co-villagers of the region are present. If at least one person from another village in the Carnian region is there, everyone switches to Friulian. It is also used under other circumstances. For example, some people who have had high school education outside of Sauris use it as the familiar conversational idiom of group solidarity. Finally, there is the German dialect of Sauris/Zahre, a unique variety of German which is limited to that town in its essential characteristics and which is used when people are talking to each other informally as fellow inhabitants of the town. In a way, you might say

that it is the nearest thing to a mother tongue Saurians have. This brief summary is, of course, an oversimplification. In any case, Norman Denison's studies of the Saurian speech community offer very interesting descriptions of sample utterances and the occasions for code switching. He is concerned with the situational factors which account for the use of one variety versus another. He identifies thirteen such situational factors and proposes a tentative hierarchy of them. His view of the immediately desirable further research is to improve on that hierarchy, by finding out when one factor outweights another, when different factors are in conflict, and so on.

Denison takes a big step toward the position I am suggesting here. He says, 'For Sauris we have to see the three languages as constituting, for community-internal purposes, one super-code, as it were. The choice of one of them as macro-textual vehicle for the linguistic component of a communicative event is like the selection in a monoglot community of register at the earliest degree of delicacy' (Denison 1968:585). He even offers some careful parallels between langague switching in Sauris and register switching within English in Britain. He does not, however, take the final step in attempting to write a unified grammar of the phonological or grammatical competence of Saurians.

As a second example, let us take a study done by Peter Trudgill, who is known to us here as a faculty member of the Institute. He has written a really excellent article about reduction and simplification in a community of Albanian speakers in Greece--a community which has been there for several centuries (Trudgill 1976-7). The members of this speech community are all bilingual in Arvanitika and Greek, and in discussing the differences between Albanian proper and Arvanitika, Trudgill notes that in some instances the slack is taken up by Greek. The simplest example is that of the lexicon; a large number of Albanian words have been replaced by Greek. Apparently an Arvanitika-speaker who does not remember a word feels free to substitute a Greek equivalent. The grammatical reductions, substitutions, and adaptations are much more complex. Arvanitika and Greek are interwoven in fascinating ways. Trudgill's concern in the article, however, is with the reduction and simplification in Albanian under the conditions of restricted use, and he makes no attempt to treat Arvanitika and Greek within a single grammatical framework. The closest he comes to this point of view is the comment '. . . the reduction to which Arvanitika is currently subjected is simply keeping pace with the extent to which Arvanites are now becoming native speakers of Greek' (Trudgill 1976-7:48).

Another example is Beatrice Lavandera and her study of the kind of language spoken by Italian immigrants in Argentina (Lavandera 1978). These Italian immigrants are bilingual. Their children typically become monolingual Spanish-speaking, but a large population of bilingual Italian immigrants persists because of continued immigration. They speak a kind of Italian among themselves, in fact several quite different kinds of Italian because they come from several parts of Italy, and they speak a kind of Spanish to the Spanish-speakers of Argentina, a 'reduced' form of Spanish, that is called Cocoliche. This variety of Spanish is very close to the surrounding Argentinian Spanish but Lavandera has shown that

it differs systematically in certain ways. In particular, the full range of stylistic variation which the native speakers of Spanish have, is lacking in Cocoliche. For example, whether you pronounce the final *s* as the sibilant, as aspiration, or as nothing at all, is a stylistic device which can be used in many ways in Argentinian Spanish. But Cocoliche-speakers either have the full *s* or they don't, as an all or none proposition. In fact native speakers of Buenos Aires Spanish use this *s* variation both word-finally and preconsonantally within the word, but Cocoliche-speakers typically drop the final *s* completely and use the full sibilant pronunciation preconsonantally. Similarly, there are different ways of using the subjunctive in indirect discourse which the Spanish-speakers have available and which the Cocoliche-speakers do not use.

In this instance the lack in Cocoliche is not interference from Italian, since the speakers use the corresponding subjunctives in their Italian. Lavandera recorded artificially elicited material in the speakers' 'purest' Italian and 'purest' Spanish, i.e., Cocoliche, ('Tell me a story in dialect.' 'Now tell me the same story in Spanish.') She also recorded spontaneous conversations in the home where the language was sometimes what the speakers later judged to be 'dialect', 'Spanish', or a mixture of the two. Her conclusion comes close to the position I am advocating here, although she does not try to write a unified grammar: 'The data indicates that for these speakers Spanish and Italian are not independent codes, at least not more so than the different registers of a monolingual speaker' (Lavandera 1978:403).

One final example, the one I find the most interesting, is a quadrilingual Indian community in Canada. Ronald Scollon recently studied a speech community in Port Chippewyan, Alberta (several articles and a monograph in press; cf. Scollon for the forthcoming) where the people have for at least a hundred years, been speaking four different languages: Chippewyan, Cree, French, and English--an Athabaskan language, an Algonquian language, a Romance language, and a Germanic language. And over the times, these languages have come to play particular roles in that community in a very complex pattern. Everybody speaks all four languages, but the phonologies and the syntaxes have become remarkable alike. For example, the phonologies tend to be like Cree and the syntaxes more like Chippewyan. But there are nice examples of patterns that cut across all four languages. Particularly interesting is a phonological variable described by Scollon. Beginning with some variation previously existing in Cree, there is now quite extensive variation between 's' sounds and '∫' sounds, including affricates of the *ts*, *t∫* type. It works this way: in general, the speakers tend to use the 's' kind of sound when they are unconsciously expressing a modernizing tendency, that is when the situation is recognized as a modernizing one or a modernizing topic is being discussed or a modernizing addressee is involved. And the '∫' kinds of sounds are associated with traditional addressees and so on. The sound variation which began in Cree has now spread throughout the system, and so we find such strange anomalies to the observer, as this example which Scollon gives: one particular woman in this speech community was heard saying [suz] for English *shoes* when teaching the kindergarten, a modernizing locale, and she was heard saying [∫a] for [sa] the word for 'beaver' in Chippewyan, when she was at home talking

with a group of elders in the community, a traditional setting. So in both cases, the pronunciation was 'wrong' in terms of what the original forms should be, but it was right in terms of the total grammar of the speech community. In this wonderful set of studies by Scollon, does he talk about a grammar of this whole system? Not at all. His first concern in one of the papers is performance aspects of traditional narratives in the society. He is trying to understand the way traditional narratives are recited, their functions in communication, how they are framed in the discourse, and so on, and he provides a very fine exposition of traditional narrative performance. In order to talk about that, he has a lot about linguistic convergence. The closest he comes to saying that it all could constitute a single grammar is this: 'The four languages at Port Chippewyan . . . have converged to a considerable extent. . . . The systematicity of discrete historical varieties is subordinated to the pragmatic needs and personal experiences of the individual' (Scollon forthcoming).

Incidentally, I would like to point out that all four of the studies have a diachronic flavor. Whether the authors intended it or not they all talked about what those languages were like before the time of description and what seems to have happened and to be happening, and they all make some cautious predictions about what might happen in the future. Also, it so happens that these fine linguists in making these descriptions, at one point or another in each one of the articles, have acknowledged reluctantly, hesitantly, and in a kind of aside, that really the linguistic repertoire is one system. What I am saying is, let's face that openly. Let's try the job of writing grammars of such complex systems. I know it is hard enough to write grammars, anyway, and what I am suggesting makes it even harder. But I still think this is the job of the linguist. As long as we keep kidding ourselves that there are no situations like this, in some way, that they are not somehow linguistically significant, then I think we are missing an important part of linguistics. We are missing an important part of how languages change through time and how language variation functions synchronically, essential features about the nature of human language. And I have a final quotation, this one from a study of mine. Back in 1957, I wrote an article called 'Two problems in Arabic Phonology,' a somewhat polemic technical description of two questions in Syrian Arabic phonology. One of the questions was the *ē/ō* versus *ay/aw* that I mentioned earlier. At the end of the article, I apparently lapsed into a kind of dreamy mood, because this is the sentence I wrote: 'I like to dream of a time when the analysis of a uniform, static, autonomous synchronic system will be regarded as merely a special limiting case in a far more general linguistics which is prepared to deal with data highly heterogenous dialectally and diachronically' (Ferguson 1957:478).

I still agree with the substance of that sentence but the addition I am making tonight is that if the variation turns out to include varieties so different that we would want to call them different languages, we might still have to put it in the grammar. In other words, I am saying that multilingualism may be a legitimate object of linguistic description.

REFERENCES

Denison, N. 1968. Sauris: A trilingual community in diatypic perspective. Man 3. 578-92.

Ferguson, C. A. 1957. Two problems in Arabic phonology. Word 13. 460-78.

------. 1978. Historical background of universals research. Universals of human language, ed. by J. H. Greenberg et al. Vol. 1, 9-31. Stanford: Stanford University Press.

------ and C. E. Debose. 1977. Simplified registers, broken language, and pidgin and creole linguistics, ed. by A. Valdman, 99-125. Bloomington: Indiana University Press.

------ and J. J. Gumperz. 1960. Introduction. Linguistic diversity in South Asia, ed. by C. A. Ferguson and J. J. Gumperz (IJAL Vol. 26, No. 3, Part IV), 1-18.

Kellogg, S. H. 1965. A grammar of the Hindi language. 3rd ed. London: Routledge and Kegan Paul.

Lavandera, B. R. 1978. The variable component in bilingual performance. International dimensions of bilingual education, ed. by J. E. Alatis, 391-409. (Georgetown University Round Table on Language and Linguistics 1978.)

Scollon, R. Forthcoming. Variable data and linguistic convergence. To appear in Language in Society 6.

Smyth, H. W. 1956. Greek grammar. Rev. ed. Cambridge, Mass.: Harvard University Press.

Trudgill, P. 1976-7. Creolization in reverse: Reduction and simplification in the Albanian dialects of Greece. Transactions of the Philogical Society 1976-7. 32-50.

MULTILINGUALISM AS A GOAL OF EDUCATIONAL POLICY

Roger W. Shuy
Georgetown University &
The Center for Applied Linguistics

In his introduction to one of the most recent collections of articles on the amazingly diverse and emotional topic, bilingualism in America, Hernan LaFontaine notes that educational policy development is difficult to articulate and even more difficult to change. In the case of the recent flurry of interest in multilingualism in this country, LaFontaine (1978:xi) observes: 'In addition to the normal inertia resisting any educational change, we have had to face political, social, and economic issues raised as barriers to the effective implementation of bilingual education programs'. Such issues are not new to education for this constitues a field which has been characterized by decades of searching mindlessly after every new method or gimmick which comes along. Now that language has surfaced as the focus of the developing policy, the inertia which so characterizes education is of immediate concern to linguists.[1]

Recent events in educational legislation are instructive about the process of change in American schooling. In January of 1974, after four years of litigation and appeals, the U.S. Supreme Court ruled in favor of a San Francisco Chinese family named Lau who claimed that the local school system had violated their constitutional right to access to education by providing that education in a language which was foreign to the learner, English. By ruling in favor of Lau, the Supreme Court said, in effect, that it is the right of every American citizen to expect the strong and the educated to deliver learning in such a way as to address the beginning points of the weak and the uneducated. Education has claimed to be doing this for years, usually referring to it as 'starting with the child where he is'. Although this principle is widely proclaimed, in reality it is clearly absent, for major American educational policy continues to follow a clearcut compensatory education model which places little value on diversity and great emphasis on making every child as much like main-stream children as possible as soon as it can be done. Lau vs. Nichols suddenly challenged all this and, with one stroke of legislation, brought legislative power to what had been almost empty verbiage in teaching and learning.

This bilingual education legislation, like much of educational legislation was actually an expression of a moral value derived from a great deal of intuition and from very little empirical research or theoretical clarity. Social scientists had not provided Congress with a clear date base for bilingual education any more than they had in the past offered ample evidence that busing would lead to better learning for the educationally disadvantaged. In fact, when confronted with the assertion that bilingual legislation preceded the knowledge upon which it could be based, the lawyers for both Lau and Aspira[2] readily admitted that such was the

case, noting further that without such legislation, the knowledge base might never be started.[3] This does not mean that no knowledge base existed prior to legislation; only that this knowledge base was far from adequate. Nor is this to cast criticism on the legislation, for it was undoubtedly well motivated and much needed. It is called to attention here rather as a humbling reminder that educational theory and research, in many instances, tends to follow the legislation rather than precede it.

In fact, the developmental model might be said to look something like the following:

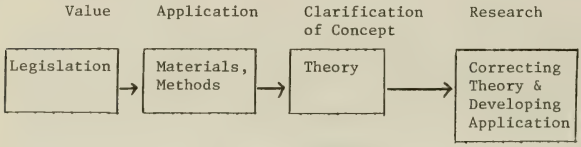


FIGURE 1

This model of development in educational programming is not at all unfamiliar in our schools. In a recent evaluation of Title I programs in America, no effort was made to assess the theoretical difference between compensatory and noncompensatory theory.[4] Children in Title I programs were not compared with non-Title I children nor was there any conscious effort to address that most critical aspect of education: the legitimacy of what the learner already knows. Instead, the evaluation assumed the legitimacy of compensatory theory and began looking for delivery system differences which mattered. The value was assumed to be correct. The materials and curriculum were developed and by the time the evaluation took place, nobody could remember that theory nor basic research had been considered. The press of the political, social and economic concerns was too great to be delayed by our not having thought about, much less, researched the issue.

The fact that such a procedure is not unusual in education is no excuse for linguists throwing up their hands in despair. Educational policy is developed no less chaotically than language policy, and we have had far too small an impact even on policy issues which bear directly on our own field.

What linguists know to be true about how language works has not been communicated to education and has had little impact on educational policy. As a field, we have been busy with our own problems of analysis and theory, and we have, with good reason, been cautious about making claims which may later turn out to be wrong. Perhaps this caution is good in some ways. Linguists should certainly try to avoid the excess claims for our field such as those which were made in the fifties, when linguistics was thought to be the savior of reading and composition. Vestiges of this overpromise remain today in the form of 'the linguistic approach' to reading (which is, in actuality, a phonological approach to decoding). Having been stung by this and the Paul Roberts composition disaster (which not only promised that generative grammar would aid writing ability but also gave an early and

somewhat inaccurate representation of generative grammar), educators stopped seeking out linguists and most linguists didn't seem to care. We went off to our own issues and let education find other easy 'believisms' to adopt.

It was not to the credit of the field of linguistics that multilinguism came into prominence with the passage of the Title VII Amendment to the Elementary and Secondary Act. Embarassing as it may be, the turn toward the direction of favoring more than one language in American education (and thus in American life) did not grow out of suggestions from the field which calls such practice its own. Since linguistics also has had little to do with stimulating other national language movements in recent years such as Vernacular Black English and the resurgence of interest in Native American languages, one might speculate about why we haven't pushed our own field, our own interest, our own expertise area. At least seven factors, individually or together, could cause linguistics to be slow in this matter.

1. *Energy.* Perhaps linguists have too much to do. We teach, write, research, and appear to work very hard but, as fields go, we are few in number. Our problems to address are endless. Some will remember when an LSA meeting seemed to have participants who believed that we were only ten years or so away from writing a complete grammar of English. The more we learned, the farther away that point seemed to get. Today there is hardly a linguist alive who feels that we are anything short of light years away from such a goal. Perhaps we do not participate in developing multilingual policy because we lack energy, but this does not seem to be the total cause.

2. *Belief.* A second possible reason why linguistics has been slow to influence educational policy is that we may not really believe in what we are doing. This is an extremely difficult suggestion to accept since there is probably no more fanatic discipline in all of academics than linguistics. Nevertheless, the question remains that if linguists practice multilingualism and if we believe in it, why is it that we have had so little influence on educational policy? Do we believe it in our heads but not in our hearts?

3. *Academic Neutrality.* With the academic stance of neutrality we may be getting closer to the reason why linguistics is slow to influence educational policy. This neutrality gets reflected in many ways. Some linguists are above getting involved with making their theory move toward practical reality. It smacks of the academic proletariat. It seems too much like getting one's hands dirty or working in the kitchen. There was a time, not too long ago, when applied linguistics was frowned upon because it cast off the stance of neutrality and entered the political, social, and economic world. The ineffectiveness of such a stance was brought home to me clearly a few years ago when a colleague and I had presented to an author cogent, academic reasons why the reading test which he had developed was in error and dangerous. His response to our linguistic argumentation was direct and simple: 'The test sells well and as long as it sells well, I'll keep it on the market.' Clearly our academic neutrality would do nothing useful for the thousands of children who were being misdiagnosed by this test. The issue was an economic one as far as the author

was concerned and the best way to address an economic issue is not with linguistic reasons but with an economic weapon such as systematic boycott.

Linguists who do not want to get involved, who just want to do their specialty and avoid advocacy have a legitimate right to do so, of course. And as a field, we can simply describe multilingualism, show how it works and develop theories about it, remaining neutral about it all. If we do this, however, we will not be doing much toward developing educational policy related to multilingualism. We will be leaving it to other people, many of whom know little about multilingualism.

4. *Vision.* A fourth reason why linguists have been slow to effect educational policy of multilingualism stems from a myopic perspective of our field. In contrast to the previously noted academic neutrality, this position is one of academic isolation which leads to a severe type of intellectual myopia. The field of bilingual education in America suffers badly from such myopia. Linguists tend to think of bilingual education as language. Social scientists think of it as culture. Lawyers consider it law. Educationists see it as materials, methods, curriculum, and teaching. All are right, of course, but it is clear that few see it from the perspective of the other.

If the lawyers for the Puerto Rican Legal Defense and Education Fund had seen bilingual education from the perspective of education and linguistics, for example, they would never have argued for a consent decree which required the New York City schools to develop, administer, and score tests in English and Spanish for over 200,000 Puerto Rican children in writing, reading, listening, and speaking to determine in which language these children could 'most effectively participate in the classroom.' Linguists could have told the lawyers that such a test does not exist and that it would take a great deal of research to discover what it means to 'effectively participate in the classroom.' They could also affirm that it is rather difficult to test for speaking ability in a paper/pencil group test. Educators could have told the lawyers that New York teachers do not look with favor on losing a week of class for such testing and that the teachers union, being generally opposed to bilingual education in the first place, would not necessarily discourage lost or falsified data.

It is difficult to imagine how this disciplinary isolation could be averted. Universities foster it and many efforts at building interdisciplinary programs have yielded products which know a little about the various disciplines but are competent in none of them. It would be far better to be fully trained in a given field such as linguistics, then work on problems in education along with educators, not in opposition to them. Team research, team teaching, and a general cooperative spirit seems to be the answer here. It is clear, however, that disciplinary isolation has slowed linguistics in its effort to influence educational policy related to multilingualism.

5. *Snobbishness.* Closely allied with the vision problem is that of snobbishness. Perhaps what we really mean is insecurity since it is

often quite difficult to tell one from the other. Disciplinary isolation seems to go hand in hand with one-ups-manship of this sort, however, and it is not difficult to see evidences of such unhappy behavior among our own numbers.

Sometimes this snobbishness is not even recognized. Clear evidences seem to be easier to locate in crimes against linguistics rather than crimes of linguists against other fields. It is no secret, however, that among the social sciences there are different research paradigms and techniques. Many experimental psychologists, for example, seem to be unaware of the legitimacy of such nonexperimental approaches as participant observation or ethnographic analysis. Recently, in fact, the NIE sponsored a workshop which addressed the conflict between quantitative and qualitative approaches to research in education. This was a very important first step since quantitative research has dominated government research funding in recent years. A story is told, in fact, of an experimental psychologist who was introduced to a classicist. The psychologist asked the classicist what he was doing and the latter responded that he was spending the year in research. The psychologist looked surprised and replied, 'I didn't know that classicists did research.'

Such snobbishness is not limited to the field of experimental psychology or to quantitative techniques. It is safe to say that linguistics is perceived as abstract, esoteric, jargon-ridden, and impractical by many school people. Linguists are often thought to be irritable, excited, unhelpful, and one-track minded. It is not difficult to see how such an impression could develop if we do no more than tell educators that they need linguistics and that, to get it, they need to take our introduction to linguistics courses. While their concerns are on reading, writing, speaking, and listening, ours are on phonology, grammar, semantics, and universals. It is easy to be thought of as snobs under such circumstances.

6. *Knowledge.* Perhaps linguistics has not affected educational policy in the area of multilingualism simply because we do not have the knowledge base with which to do it. We know some things about first language acquisition (at least up to age ten or so) and a few things about code switching and second language learning. But we are practically helpless in answering questions of the schools and the courts which concern fluency (what does fluency mean anyhow?), the effect of cognitive growth and language switching, the exact dimensions of literacy and the language of literacy. Even more embarassing, we do not have answers to the educators' questions on the most basic issues such as what really matters most in learning a second language. Does a nonnative really have to learn to speak a new language like a native? Won't the nonnative lose something important if he speaks the target language the way native speakers do? These and many other potential research questions are unanswered. The courts believe that someone has the answers to these questions. Linguists clearly do not. When asked, we admit it. Then the schools have to act as though the answer were clear. Since linguists haven't given the answer, the schools muddle along as best they can. They make policy without us because someone has to do it.

7. *Internal Controversy.* However much linguists might consider internal controversy to be a healthy thing which enables us to grow and expand, we are often perceived by educators and others to be constantly argumentative and undecided. As we all know, linguistics is a field in which people engaged in bitter public disagreement can turn around and, unexplainably, go to lunch together as friends. This is not always the case, of course, but it is clear that others do not always perceive us in the same way that we may perceive ourselves. In addition, the fact that our field seems to have theoretical revolutions with great regularity leaves other fields in a quandry about how to relate to us.

Of these seven factors which one might suggest to explain why linguistics has been slow in affecting educational policy with respect to multilingualism, five stand out. It is easy to reject the suggestions that we lack energy or belief. We can probably be faulted when we attempt to be academically neutral. We have undoubtedly been myopic about multilingualism, seeing it only as a linguistic phenomenon. Accurately or not, we have been perceived as snobbish and prone to excessive internal controversy. It is the factor of our inadequate knowledge base which should concern us the most, however. Multilingualism is clearly within our territory. Linguistics may not own the field but it can hardly be said to be marginal to it. That we have so few answers to offer the courts or the schools as proof of what works or even what is, is no credit to our discipline.

In general, linguistics has not taken an active role in trying to initiate language policy in this country. We have not anticipated what is coming next. We have not analyzed our market. We have not tried to forecast. Indeed, linguistic involvement in language issues has tended to be responsive rather than interactive or formulative. We have let other forces of society set the stage and make mistakes. Then we come on the scene and complain bitterly about how stupid the schools are to have set their current courses.

Our involvement with educational issues relating to Black English Vernacular in the sixties is a case in point. School problems of Black children were recognized first in the schools by people like Ruth Golden in Detroit. It is not possible at this time to know whether or not Ruth Golden asked linguists for help when she attempted to develop standard English teaching procedures in that city. In a way it is immaterial for, if she had asked, she would have found little of use to her. So she went ahead on her own. She developed the Golden Tapes, a set of audio tapes which addressed certain Black English pronunciations. Mrs. Golden found it particularly troublesome that students at Central High School pronounced it Cintral High School. So she developed tape drills to correct this. The point here is not that Ruth Golden selected a poor feature to focus on. It is, rather, that there was no linguistic help for her to rely on. She had to go it alone. Five years later, by 1968, linguists could tell her that the /I E/ collapse before nasals does not really matter. Other features count considerably more heavily. Why couldn't linguists have anticipated the Black English development and geared up for it in advance? We had to be responsive rather than interactive or formulative. What is worse, we were not even anticipatory.

Other fields have developed an anticipatory capacity with some degree of reliability. We are all familiar with difficulties involved in weather-forecasting and economic predictions but even something as prosaic (is this snobbishness again?) as sales forecasting may be able to offer insights. Business is not merely responsive; it interacts to change societal conditions to its own liking through advertising, lobbying, bribing, three-martini lunches, and many other ways. Lest there be any confusion on this matter, let me state clearly that I am not suggesting that linguists begin advertising or bribing, or having three martini lunches, but I think it might be useful for us to consider two of the common processes of sales forecasting.

One process simply looks at the past and notes recurring cycles and locates the current situation somewhere in that cycle. This is a bit like projecting language change based on a developmental line from as far back in time as we can go to the present then extended (perhaps a dotted line) into the future, based on past development.

The second process of sales forecasting is the careful analysis of market conditions. Sales forecasters predict future conditions by matching controllable and noncontrollable factors. The controllable factors are things like what we know, how much equipment we have, our financial condition, our personnel, and our reputation and image. The noncontrollables are the social forces at work, demographic factors, ethical forces, political forces, cultural environment, international issues, and the development of new technology. For example, if technology should suddenly create a way to make congenitally deaf people hear normally, the political issues involved in the signing (ASL) vs. oralists camps of deaf education would suddenly seem less critical. Without such technology, however, linguists have been stoking up the controllable resources to try to make oralists understand that signing is the native language of deaf people.[5]

The point here is that linguists will continue to find it difficult to affect educational policy involving multilingualism as long as the field is merely responsive to developments created outside of linguistics. In order to be anything other than responsive, linguistics must develop ways of anticipating societal needs in advance and gearing up for them before they happen. The rest of this paper will provide five suggestions about what linguistics must do to become more active in developing educational policy involving multilingualism.

1. *Become accepted by education groups.* Work with educators (in team, if possible). If linguistics is to be recognized as an ally of education, we will need linguists who can communicate with educators. We should be writing in their journals, speaking at their conventions, and serving on their panels. If one wants to do something useful about doctor-patient communication, one says it where it will be heard by the right people. One writes for a medical journal or speaks at a medical convention. If we reject the idea of speaking at an education convention on the grounds that such conventions are of low calibre, we are either guilty of the snobbishness discussed earlier or we have picked

the wrong convention. Two very important considerations must be kept in mind when one engages in this sort of activity. First, there is a pecking order in education just as there is in linguistics. Failure to cite the perceived acceptable literature can limit outsiders' positive reactions considerably. Secondly, educators will be interested in their own problems primarily and it is up to us to show them how linguistics can contribute to their problems. Administrators face the problem of how to assign pupils to bilingual programs or regular programs. They frequently rely on tests, such as language dominance tests, to make such decisions. Linguists can argue effectively that language dominance is a peculiar concept in search of a definition but this will not make the administrator's problem go away. Eventually we will need to do more than debunk current practice. We will need to offer a viable alternative to the administrator's problem. Writing in education journals, appearing at education conventions, and serving on education boards provide the evidence of our concern. Such evidence will enable us to be accepted by the education groups through which policy on multilingualism will likely be made.

2. *Get involved with the power structure.* Earlier I commented that lobbying might not be appropriate for all people. Nevertheless the business of influencing educational policy requires a broker of some sort. It is possible to get oneself appointed to textbook selection committees, to write to the NIE Director, to read proposals, or to sit on evaluation panels. Those of us linguists who read proposals in areas related to language in education, speech, or humanities realize full well that our job is not only to evaluate a proposal but also to represent linguistics to our fellow panelists (who are most likely to be psychologists) and to actually teach them about research designs, strategies, and knowledge from our field.

3. *When possible, get involved with advocacy.* To make multilingualism become education policy it would appear that linguists will need to become considerably more activist than we ever were before. A linguist who is accepted by education and who is involved in the power structure stands a better chance of being asked to serve as an expert witness in a law suit involving bilingual education than one who is not so involved. The decision to be an advocate in such circumstances opens one to considerable complication, however, as the following incident will indicate.

In December of 1976, the New York City School System had just entered into a consent decree with a New York Puerto Rican rights group, Aspira of New York. As noted earlier, the consent decree specified that the school system test to determine in which language, Spanish or English, the over 200,000 New York Puerto Rican children could most effectively participate in the classroom. The test was to be developed, administered, and scored by the following June. It was to measure ability in both languages in reading, writing, listening, and speaking. The department of research in the schools called on the Center For Applied Linguistics to develop

the test. One quick look at these impossible requirements prompted CAL to suggest a different strategy. Rather than to do a job which had to be based on an as yet unidentified and unresearched knowledge base, CAL suggested that New York throw itself at the mercy of the federal judge and tell the truth: that nobody really knows how to do this as yet and that it would take a year or two of solid research and development to be even near ready to carry out the task.

CAL was not effective in convincing the director of research, however, and New York went at it on their own. Approximately two weeks later, the test was completed and the director of research happily claimed that he could do in two weeks what CAL projected would take a year or so.

At this point, the Puerto Rican Legal Defense and Education Fund, legal council for Aspira, called CAL to request help in determining how good this new test, the Language Aptitude Battery, actually was. CAL took great delight in this request for it gave linguistics a golden opportunity to address the issue of bilingual placement. Part of the work was to explain to the judge, in court, the weaknesses of a test which claimed that discrete point items such as the /š/ vs. /č/ contrast in words like *shoes* and *choose* are the critical measurements of language ability. Another task was to explain to the judge that a test written in English would not be equally valid when translated into Spanish. In addition to these validity issues, CAL pointed out to the court many of the ambiguities in the wording of both the questions and the instructions.

A second incident involving CAL advocacy work took place in Chicago where, in 1976, the Office of Civil Rights declared that the city schools were out of compliance with the Equal Rights Amendment, largely in the area of minority language speakers. The procedure for identifying bilingual children was again the case. Chicago represented itself at the crucial hearing as identifying children in need of bilingual education by administering a fifteen minute video tape to Chicago teachers. They observed three minutes of five different children of varying English language ability. Each child was said to represent one grade on a five point scale of English competence. The teachers were to see this performance, then use it as a model and guide for classifying all of their limited English speakers. Again it was not very difficult to write an affidavit describing what was wrong with this procedure. The complex part, however, was that the director of bilingual education in that school system is an extremely competent person who sincerely wanted to improve the situation. To make matters worse, the attorneys for the Chicago schools did not put Chicago's best foot forward. Chicago was doing better things than this court performance showed. The business in advocacy roles often puts the linguist into a very touchy situation. Should we continue to help OCR attack what Chicago has been doing even though we know that they are in competent hands and making steady progress? We

gambled to go ahead on the assumption that by exposing the school's presentation we might ultimately strengthen the program and the director's hand at the same time. As it turns out, this is exactly how it worked. OCR won the case and the Chicago schools were threatened with having all federal funding withheld. There was a quick change in strategy shortly after, however, and the program became stronger than ever.

Once one enters into policy development such as this, however, one runs certain risks. I was personally frightened that my good relationship with the director of bilingual education might deteriorate because of the court case. I gambled that it would not and, fortunately, I was right.

Getting involved with advocacy in this way is something which few academics ever think they will do when they first get seduced by their field. My own personal dreams of living a tweedy, halls of ivy existence vanished like vapor as soon as I realized that if I really believed in what I was teaching, I might also be expected to do something about it. This doing can be interpreted as talking or teaching, at least for a while--or maybe even writing papers ang giving speeches. But eventually these things come to seem oddly preliminary and removed from the real arena. At such times one does not stop being a linguist and start being a salesperson. Instead one continues to be a linguist and tries to determine ways to influence policy. If the linguist does not do this, others who know our field less well most certainly will. And this is exactly what has happened far too often in the past. The friends of linguistics, those who had a little knowledge, are the ones who did us in, who promised the moon and who saw linguistics as the answer to education's problems. Obviously that tack has not worked. Educational policy is no better for it and it has certainly not been enlightened by it. Perhaps it is time to try it from the obverse side of the coin. If educators do not understand language well enough to lead in the areas of educational policy related to language, perhaps linguists can develop strategies toward understanding education well enough to do it. To do this there is probably no better subject matter than multilingualism. And there is no quicker way to find oneself in the middle of the policy issue than to become involved with an honest-to-goodness law suit involving multilingualism.

4. *Research needed problems.* A fourth way for linguists to become involved in the educational policy of multilingualism is to direct our research toward the needed problems. At this time we still do not know what matters most and what matters least in learning a second language. We need a much broader and deeper research base for bilingual education than is now available. To date, research in bilingual education has been considered more an education issue than a linguistic one. The tendency of the research has been to compare one methodology with another or one curriculum with another. However, we have now reached a period in the development of the field when people are demanding evidence of success or failure of

the AIR evaluation of Title VII programs is that it did not understand the basic principles of language variability.[6] It is difficult to put faith in any assessment of the effectiveness of bilingual education which does not distinguish between acceptable and unacceptable variation and which does not address language variability as it manifests itself in the dimensions of social, regional, age, sex, race, education, occupation, intention, genre, topic, etc. It is also growing increasingly difficult to accept the discrete-point criteria used over and over again as evidence for language ability in either language.

Recent research in linguistics has focused more and more on the importance of core syntactic, semantic, and pragmatic processes and less and less the isolable phonological, morphological, and lexical items. It has focused on the holistic more than on the segmental skills. Bilingual education programs need to focus more on this holistic view and more the sort of variability involved in asking questions, making requests, seeking clarification, taking a conversational turn, and reporting what others have said, for example, than on pronunciation accuracy or inflectional perfection. This direction is not merely a tag-on to existing programs; it is the starting point for language learning.

Any effective research on language in bilingual education programs will need to have, at its base, an adequate linguistic theory of language learning. Whether by accident or by design, there is, at the core of bilingual education, a satisfying theory of language learning. From research on first or native language learning we know that infants communicate ideas first, then later they perfect their phonology, lexicon, and grammar. Traditional language programs cause second language learners to try to reverse this process, getting the phonology, lexicon and grammar first and hoping, later, for the development of language use. When subject matter is taught in a second language, the focus gets put right back on the idea of subject again, just as it is for infants. The resulting learner language is, in comparison to that of a mature speaker, rather sophisticated, just as it is for infants. But it is the process that is sophisticated and thousands of years of experience have shown it to be the superior method. Thus, bilingual education, for this very reason, has at its disposal an adequate language learning theory: a deep to surface structure sequence which has been proved successful by millions of infants as they learn their native tongues. Such a theory contrasts sharply with the surface to deep structure sequence found in most second language classrooms.

With the advantage of this deep to surface theory, bilingual education research could easily capitalize on the advances made recently by other first language learning research. We know, for example, that pronominalization is learned early and easily by children whereas preposed adverbials are learned rather late. Nancy Yanofsky (1978) has recently shown, however, that reading instruction does not take advantage of this knowledge. A number of texts analyzed in her research show that preposed adverbs are

frequent in beginning reading instruction while pronouns are introduced rather slowly.

The areas for research into language variation and bilingualism are many. Careful descriptions of the production of such variability need to be made in order to determine how the speakers actually talk. Likewise, sociolinguistic attitude studies need to be made in order to determine community values toward the various language varieties used in school programs. We desperately need some clarity in the area of language assessment, particularly as it reflects a learner's sequential progress in a number of school and non-school settings. We need a dose of realism in the development of curricula in language teaching as well, with a focus on the contributions which linguistics can make in bringing about a radical restructuring of language learning classrooms by focusing on core syntactic, semantic and pragmatic units of language rather than the minute, phonological, morphological, and lexical aspects.

5. *Become well-placed.* A fifth way for linguists to influence the educational policy of bilingualism is to get ourselves well placed. One of the few benefits of the depressing job market in linguistics today is that it is forcing us to think of working outside of academia. And where is educational policy often forged? Outside of academia. It will do our field, and the subject of multilingualism a great deal of good if folks who know about language are placed in government agencies in which language issues are directly or indirectly addressed. Likewise it may do multilingualism good to have linguists working with speech therapists, helping to sort out speech problems from multilingual differences and as lawyers, litigating cases involving language issues. It may do multilingualism good to have linguists working in the advertising agencies in New York, monitoring language issues on television ads and it will most certainly do the field of multilingualism good to have linguists working for publishers. All of these positions, however promising they may be for influencing educational policy, require a sacrifice. Marcia Whiteman has had to learn to be a bureaucrat in order to exercise her linguistic judgment on language issues at N.I.E. Andy Garfinckel has had to learn the advertising business as a trainee in order to work in television advertising. George Williams, Mary Levy, and others have actually gone to law school in order to be able to better apply their language knowledge toward policy issues. Faye Vaughn-Cooke, a linguist, works in speech therapy saving minority speakers from being designated as having severe pathologies. And as of September we will have, as far as I know, our first Ph.D. in linguistics enter Georgetown University's medical school where Johnnie Smelz will get the necessary medical credential to serve her consuming interest in neurolinguistics. To become well-placed will require sacrifices of a sort which linguists have not had to make in the past. It may be easy for me to say, since I don't have to make such a sacrifice, but I think our field will be far better for it in the long run.

By carrying out these five suggestions: becoming accepted by educators, getting involved in the power structure, becoming advocates,

researching the needed areas, and getting ourselves well placed, linguistics can improve its anticipatory capability. We need to do more than to forecast the next language-related thrust, for there is a multitude of language-related issues in education already before us. Almost all aspects of testing (multilingual or monolingual), language learning, reading, and writing are deeply involved with our field. The issue of multilingualism, however, is perceived by the public as closer to linguistics than these other matters and for this reason, multilingualism is a prime candidate for a strong push from our field. Taking our cue from sales forecasting, we can most certainly address the controllable issues: what we know and need to know, our equipment, materials, personnel, and reputation. But we can also give considerably more attention to the uncontrollables: economic factors, political developments, demographic changes, ethical forces, and new technology. Obvious economic factors may suggest that we gear up for knowing Arabic. Demographic changes may suggest that we gear up for teaching Samoan in Los Angeles. As social changes go, we addressed issues of language and sex a bit faster than we got to language and race. Linguists got a late and rather disastrous start in the bilingualism push and we soon found ourselves thought to be marginal to the education departments of the country. Our inability to affect educational policy is a direct result of our failure to anticipate the movement before it happened and our lethargy in gearing up after it did.

Worse than that, linguistics has not articulated some of its own most basic premises in such a way that the education world can see and react to. It is patently obvious to anyone who has learned language, for example, that there is no way in God's earth to try to speak a new language without making mistakes in it. Progress can sometimes be noted, in fact, by the increasingly sophisticated kinds of mistakes which learners make. Years ago Labov referred to this concept as 'stages in the acquisition of Standard English.' All linguists know that children learn their native language in much the same fashion. Somehow this concept has not come through clearly to educators, who seem to feel that speakers should be assessed on some sort of right-wrong binary scale.

Likewise, linguists have not well articulated their own basic premise that variation in language is a healthy, necessary, human, and beautiful thing.

In addition, in bilingual education programs, language variability will need to be displayed as a resource rather than as a handicap. Educators, monolingual or bilingual, have been very slow to grasp this principle. Journalists seem to accept it, at least partially, when they argue that a good writer must write one way for one audience and a different way for another. Poets have also seen the principle or they would never be able to vary, to twist or to torture language in a poetic fashion. The openness of poets to this principle is rather apparent in a poem about bilingualism written a few years ago by Ogden Nash:

> There are countries where being multilingual is of no
> advantage to the tourist,
> Because no matter how impeccable his grammar and idiom the

natives refuse to understand him unless his accent and inflection are that of the purist
They tell of an American diplomat who lost his suavity never to be restored
When a Paris taxi driver denied the very existence of his destination because he pronounced Place de la Con-corde instead of Place de la Con-corde.
(At this point I must for the justly maligned Manhattan hackers interpolate a kindly word;
None of them to my knowledge has ever rebuffed a Parisian asking to be taken to the corner of Sird Avenue and Sirty-sird.) (Nash 1971)

Whether or not Nash is right about the contrastive attitudes of the Americans and French concerning language variability, he has captured something about what we need to know about bilingualism if we are to affect educational policy in any significant way. And there is a great deal more.

NOTES

[1]One might add, however, that if linguists were on top of the situation, issues of reading, language arts, testing, and equity would also be our concerns.

[2]The Aspira Consent Decree is a similar piece of legislation relating to New York City Puerto Rican children.

[3]These assertions were made at the TESOL meeting in Los Angeles in March, 1975.

[4]Instructional Dimensions Study.

[5]As a matter of precedent for policy, Georgetown University has recently accepted English as satisfying the foreign language requirement of an American congenitally deaf student. Likewise Georgetown has accepted Sign Language as meeting the foreign language requirement of a native English speaker.

[6]It must be noted, however, that there are many other reasons to criticize this evaluation, some of which have to do with design and instrumentation, regardless of language unsophistication.

REFERENCES

LaFontaine, Hernan (ed.). 1978. Preface. Bilingual Education, p. xi. Wayne, N.J.:Avery Publishing.

Nash, Ogden. 1971. Speak to me only with thine eyes, and please, not too fast. Look. Oct. 5.
Yanofsky, Nancy. 1978. Pronominalization and adverb preposing in reading texts. Unpublished ms. Georgetown University.

FORMAL VS. FUNCTIONAL CONSIDERATIONS IN PHONOLOGY*

Morris Halle
Massachusetts Institute of Technology

In his highly instructive paper 'Functional Explanations in Generative Phonology' Michael Kenstowicz (to appear) distinguishes two kinds of explanation in phonology, 'internal' and 'external'. According to Kenstowicz, 'an "internal" explanation attributes the presence of some (part of a) rule or other aspect of phonological structure to another independently motivated portion of the grammar, while an "external" explanation seeks motivation in terms of something that exists outside of and independently from the grammar.' Appeals to simplicity, to feature counting, or to the fact that certain rules are cyclically ordered would be examples of 'internal' explanations for certain phenomena, whereas appeals to ease of articulation or to avoidance of ambiguity however defined might be viewed as examples of 'external' explanations. It is the latter kind of explanations that traditionally have been labelled 'functional'. In what follows I examine a number of such 'external' explanations and attempt to show that in each case an alternative 'internal' explanation is not only available, but also preferable. While I do not wish to be understood as claiming that these examples prove that the search for 'external' explanations should be abandoned forthwith, the discussion below does cast doubt on the claims that have been advanced by proponents of 'functional' explanations.

It is an obvious fact that many languages have rules that neutralize contrasts between distinct lexical items or morphemes. Every beginning student has no doubt been told several times that Russian, German, and a host of other languages have a rule that devoices word-final obstruents. As a result, hearing the utterance [žénskij rɔ́t], a Russian does not know whether the speaker is referring to a female's mouth (/rɔt/) or to feminine gender (/rɔd/). This ambiguity might in principle present difficulties for communication, and languages might, therefore, be expected to avoid rules that result in the neutralization of distinctive contrasts. As a matter of fact, neutralization rules are attested in practically every language and the phonetic ambiguities which they produce do not appear to have great practical effects. It might, therefore, appear that the ambiguity resulting from neutralization of contrasts plays no role whatever in phonology.

This conclusion, however, does not seem altogether justified in the light of interesting observations made by Jonathan Kaye (1975) that degree of ambiguity (or its complement 'ease of recoverability') can explain certain otherwise quite puzzling phonological phenomena. I discuss here the simpler of the two examples analyzed by Kaye, that of Maori. In this language word final consonants were deleted in the alternations illustrated in (1):

(1)	Verb	Passive	Verb	Passive	
	awhi	*awhitia*	*mau*	*mauria*	
	aru	*arumia*	*patu*	*patua*	etc.

Kaye observes that 'since the consonants showing up in the passive forms are not predictable, the standard phonological analysis would posit underlying consonants for the verb stems which are deleted when not followed by a suffix' (pp. 248-49). It has, however, been argued by Hale (1973) that this is not a correct analysis. Instead Hale proposes that in Maori all lexical morphemes end with a vowel and that the passive morpheme has eight allomorphs whose distribution is governed by an idiosyncratic property of the stem.

Kaye doubts the explanation advanced by Hale that an 'underlying form in Maori with final consonant would violate a surface canonical constraint against final consonants . . .' because ' many languages, e.g., Karok, Shawnee, Arapaho, Atsina, have rules that insert a prothetic consonant (typically ? or *h*) in front of (underlying) vowel initial words. In such languages no word begins with a vowel on the surface, yet . . . there is no evidence of restructuring of the vowel initial words in any of these languages. A theory based on surface canonical forms cannot distinguish between the Maori case and the Karok case. Having vowel initial stems violates the surface constraints of Karok just as much as having consonant final stems does in Maori' (p. 249). What distinguishes the Karok case from that of Maori according to Kaye is that the consonant deletion rule of Maori produces multiply ambiguous strings, whereas the glide insertion rule of Karok does not affect the ambiguity of lexical entries. Hence in Karok no restructuring is evident, whereas in Maori a very similar change resulted in restructuring.

I am not fully persuaded by Kaye's account. It seems to me that the ambiguity of a word such as /awhi/ is unchanged by the reanalysis. Whereas on the old account the word /awhi/ is multiply ambiguous--at least, in principle--by virtue of the fact that it could arise by word final consonant deletion from /awhit/ as well as from /awhim/, /awhir/, awhik/, etc., /awhi/ is equally ambiguous on the new account because now the lexicon can include--at least, in principle--many items represented as /awhi/ which are distinguished one from another by the different passive suffixes that they require; /-mia/, /-ria/, /-kia/, etc.

There is, however, an alternative account available. Hale has pointed out to me in a recent conversation that the passive and the gerundive are the only forms in Maori where consonant alternations of the type illustrated in (1) are to be found. If we attempt to capture this fact with the help of a word final consonant deletion rule we are in effect implying that such alternations are the rule, rather than the narrowly circumscribed special cases that they in fact are. Moreover, this solution fails to capture yet another regularity of the language; namely it implies that morphemes can end either in vowels or in consonants; whereas if the passive and gerundive suffix are reanalyzed as having consonant initial allomorphs, all morphemes of the language end with a vowel. Thus, reanalysis does two things at once: it eliminates the word final consonant deletion rule from the phonological component and it restricts the canonical form of morphemes

in the language. Arrayed against this is, of course, the additional cost of a complicated and ad hoc rule governing the distribution of the two suffixes. Although there does not exist at present a proposal of how to convert these considerations into a mechanical evaluation procedure, such an evaluation is surely not unimaginable and has implicitly been utilized quite widely. Consider, for example, the treatment of the verb *to be* in phonological descriptions of English. It is invariably treated as totally exceptional and attempts are not made to integrate it into the phonological and morphological rules of the language. Whatever simplicity considerations underlie this decision--which to most linguists is so uncontroversial as not even to merit serious discussion--will also correctly handle the case of the Maori passive.

Thus, it would appear that the Maori case cannot be taken as evidence that functional explanations are to be sought in preference to formal ones. While this claim was not explicitly advanced by Kaye it is specifically made by Zimmer (1975) who draws attention to very interesting facts in Turkish phonology which according to him 'highlight a kind of problem that the old simplicity measure signally failed to address itself to, and that is not solved by most recent proposals' (p. 556). I shall attempt to show here that Zimmer is mistaken in this assertion, that the old simplicity measure does not only address itself to this problem, but in fact forces the solution that Zimmer regards as the correct one. The facts of interest are as follows. Turkish has a rule devoicing stops and affricates in syllable final position:

(2) a.	absolute	plural	objective	
	ip	*ip-ler*	*ip-i*	'rope'
	dip	*dip-ler*	*dib-i*	'bottom'
	at	*at-lar*	*at-ɨ*	'horse'
	kanat	*kanat-lar*	*kanad-ɨ*	'wing'
	sač	*sač-lar*	*saǰ-ɨ*	'hair'
	aač	*aač-lar*	*aaǰ-ɨ*	'tree'

In the Istanbul dialect certain stem-final /k/'s alternate with Ø when followed by a vowel, while others do not:

(2) b.				
i)	*ayak*	*ayak-lar*	*aya-ɨ*	'foot'
	inek	*inek-ler*	*ine-i*	'cow'
ii)	*ok*	*ok-lar*	*ok-u*	'arrow'
	kök	*kök-ler*	*kök-ü*	'root'

Historically, these alternations between /k/ and Ø are due to the fact that in the Istanbul dialect /g/ was deleted intervocalically. A possible account for the facts--and one adopted by Lees in his 1961 monograph--is basically identical with the historical development. The grammar includes a rule ordered after the syllable final devoicing rule which deletes intervocalic /g/:

(2) c. $\begin{bmatrix} \text{-sonorant} \\ \text{-coronal} \\ \text{-labial} \\ \text{+voice} \end{bmatrix} \rightarrow \emptyset \; / \; [+syl] \_\_\_\_ + [+syl]$

It will be noted that this analysis is abstract in the sense that it requires underlying representations in which a distinction is made between /k/'s that undergo rule (2c)--these are represented as /g/,--and /k/'s that do not undergo rule (2c)--these are represented as /k/.

This analysis of Lees' is contested by Zimmer who argues that an observationally adequate description of the Istanbul dialect requires instead of (2c) the rule (2d), which deletes intervocalic velar obstruents in the third and subsequent syllables of the word:

(2) d. [-son, -cor, -lab] → ∅ / [+syl][-syl]$_0$[+syl]___[+syl]

In particular, Zimmer cites such Western loan words as:

(2) e. *bek beki* 'back'
kartotek kartote-i 'cardfile'

and refers to experiments with nonsense words all of which show that deletion depends on position in the word rather than on underlying phonological representation. Zimmer views this result as evidence in favor of functional over formal considerations, and he is understood in this way also by Kenstowicz, who writes: 'Turkish speakers have thus opted for a more concrete analysis of *k*-deletion. The reason appears to be that from the point of view of the language learner, it is easier to distinguish the two different kinds of *k*-final morphemes on the basis of an observable property (the number of syllables) rather than on the basis of a property (underlying /g/) whose existence follows from rather sophisticated reasoning.'

I do not believe that this conclusion is justified. What Zimmer has shown is not that Actual Speakers, unlike Ideal Speakers or linguists, are unable or unwilling to engage in 'rather sophisticated reasoning'. Instead, Zimmer has shown that the solution proposed by Lees is wrong, simply because it fails to characterize the data, especially the data in (2e). It hardly needs to be said that from an incorrect rule no theoretical conclusions whatever can be drawn.

In his paper Zimmer observes that the abstract analysis (one including rule (2c)) can be made to work by adding a morpheme structure condition that 'post-vocalic morpheme-final /k/ occurs only in monosyllabic morphemes, and post-vocalic morpheme-final /g/ only in polysyllabic ones' (p. 561). This condition, which must explicitly be added to the grammar, can be formally expressed as shown in (2f):

(2) f. $\begin{bmatrix}-\text{son}\\-\text{cor}\\-\text{lab}\end{bmatrix} \rightarrow \begin{cases}[-\text{voice}] \;/\; \#\ ([-\text{syl}])\ [+\text{syl}]\ \_\_\_\ +\\ [+\text{voice}] \;/\; [+\text{syl}]\ [-\text{syl}]_0[+\text{syl}]\ \_\_\_\ +\end{cases}$

In other words, Lees' solution can be made observationally adequate if in addition to rule (2c) it contains also the morpheme structure condition (2f). It is this adequate solution which must be compared with Zimmer's alternative; i.e. a solution in which deletion takes place in accordance with rule (2d). The latter solution will also require a morpheme structure

condition which, however, will state a simpler fact than (2f), namely that all morpheme final postvocalic velars are voiceless. We represent this condition formally as in (2g):

(2) g. [-son, -cor, -lab] → [-voice] / [+syl] ___ +

We thus have two alternative descriptions that cover the same range of facts: one including the rule (2c) and the morpheme structure condition (2f); and the other composed of rule (2d) and morpheme structure condition (2g). We observe that while rule (2c) is simpler than (2d) the difference amounts to just a single feature. On the other hand, morpheme structure condition (2f) is considerably more complex than its counterpart (2g). Thus, on purely formal grounds the solution advanced by Zimmer must be chosen over the empirically adequate version of the solution à la Lees. In other words, the issue is not between a solution requiring no sophisticated thought and one that demands elaborate reasoning, but rather between two solutions that require the same amount of sophistication, where one is somewhat simpler than the other; simplicity being measured by the old feature-counting metric of SPE. I conclude that the Turkish example provides no evidence whatever for the proposition that purely formal considerations may be overridden by considerations of a functional sort.

It is important to be clear about the fact that simplicity considerations are relevant only in a restricted class of cases--i.e. in cases like the Turkish velar deletion just discussed where the same theory admits two or more alternative formulations that cover the exact same body of facts. Appeal to the simplicity metric is irrelevant where alternative descriptions either cover different sets of facts or where they employ different theoretical devices. Thus, appeals to simplicity--in the narrow technical sense of the term employed here--cannot decide, for example, between a description that admits reference to metrical trees of the kind exemplified in Liberman and Prince's (1977) recent article as against one based on a theory like that of Chomsky and Halle (1968) that has no place for such theoretical constructs. The fact that simplicity considerations are irrelevant in such cases does not mean that there are no rational ways of choosing between alternative theories but only that this choice has to be made on grounds other than simplicity.

In several papers which deservedly have attracted wide attention Paul Kiparsky (1973) has suggested that certain limitations be imposed on the characterization of phonological facts. In the earliest of these papers 'How Abstract is Phonology?' written in 1968, Kiparsky proposed that phonological theory should be so constrained as to rule out 'the diacritic use of phonological features and the phonologic use of diacritic features.' The latter use, according to Kiparsky, is exemplified by rules that 'have the form of phonological rules but operate on diacritic features' (p. 16), and I have nothing further to way here about rules of this kind. The diacritic use of phonological features, on the other hand, is illustrated by cases where two forms are given different underlying representations 'for the sole purpose of classifying segments into those that do and those that do not meet the structural description of a rule' (p. 15). Such rules--of which the /g/-deletion rule (2c) in Lees' account of Turkish is

a typical example--result in what Kiparsky terms 'absolute neutralization', and in his 1968 paper Kiparsky proposed to exclude rules of absolute neutralization by imposing the so-called alternation condition which requires 'that morphemes which are always phonetically identical must have the same underlying representations' (p. 18).

Kiparsky's condition is said to render grammars more easily learned; since the relationship between underlying and surface representations is more direct, the learning of the underlying representations requires much less of that 'rather sophisticated reasoning' which many linguists appear to be uneasy about attributing to the average man or woman on the street. I should like to remark that while Kiparsky's condition renders grammars less abstract and hence, perhaps, more easily learned, it does not speak directly to the question of what role 'formal' vs. 'functional' considerations play in the functioning and development of languages. Kiparsky's proposal is a proposal about the form of underlying representation of morphemes and is thus on a par with every other condition imposed on the form of grammars. What differentiates it from some of those--e.g. from the conventions on rule ordering--is that we have difficulty in thinking of a functional motivation for the latter, but can readily think of one for the alternation condition.

As is well known, Kiparsky's paper gave rise to an extended discussion, and as a result of this discussion as well as of additional independent work, Kiparsky has modified his views to a considerable extent. The modified position is outlined in the paper 'Abstractness, Opacity and Global Rules' (1973). Its most important feature is that whereas previously 'the Alternation Condition (had) been formulated . . . as a limitation on underlying forms,' the new proposal reformulates it 'as a limitation on the application of phonological rules' (p. 65). Specifically, Kiparsky proposes the requirement (3):

(3) Nonautomatic neutralization processes apply only to derived forms.

One of the most important facts that led Kiparsky to revise the Alternation Condition in the above manner was his discovery that there exists a class of examples where absolute neutralization is blocked in certain contexts and admitted in others. In particular, rules of this class apply only in 'derived' contexts; i.e. either across morpheme boundary, or morpheme-internally, if a crucial part of the context triggering the rule is itself derived by an earlier rule. They do not apply morpheme-internally where the triggering context is present already in the underlying representation. I regard the discovery of the existence of rules of this kind as one of the most significant empirical finds in modern phonology.

Kiparsky illustrates how such rules operate with the help of the following two rules of Finnish phonology:

(4) a. *e* → *i* / ___ #
    b. *t* → *s* / ___ *i*
    c. i) /halut+i/ → /halus+i/ 'wanted'
                (4b)

ii) /vete/ → /veti/ → /vesi/ 'water'
(4a)
iii) /tila/ 'place'
/itikka/ 'mosquito'
/neiti/ 'Miss'

It should be noted here that in Kiparsky's formulation restriction (3) does not apply to all phonological rules but only to phonological rules which are neutralizing and which, moreover, are nonautomatic, i.e. which admit specific exceptions. In their recent book on phonological theory, Kenstowicz and Kisseberth (1977) raise cogent questions about this restriction: 'Why should automatic neutralization rules apply in both derived and nonderived contexts, but nonautomatic neutralization rules just in derived contexts? . . . Suppose that there are one or two items that are exceptions to [a] rule of neutralization . . . Why should the existence of such random exceptions have the consequence that "an alternation . . . cannot be accounted for" by appealing to a rule of absolute neutralization' (p. 212). As a matter of fact, Kenstowicz and Kisseberth adduce two examples where a nonautomatic neutralization rule must be permitted to apply in nonderived contexts. I briefly sketch one of these examples. In the Yawelmani dialect of Yokuts vowels are shortened in closed syllables. This shortening rule is a neutralizing process since it results in the merger of long vowels into their short counterparts, which otherwise are distinct phonemes in the language. The shortening rule is also nonautomatic: it does not apply before the causative-repetitive suffix /lsa/. In view of this, restriction (3) should block the application of the shortening rule in nonderived contexts. This, however, is not the case. Kenstowicz and Kisseberth cite the form [?aml + al] 'might help' which must derive from an underlying [?a:ml + al]. Thus shortening, which is a nonautomatic neutralizing rule, must be allowed to apply in a nonderived context, in violation of (3).[1] Moreover, in his dissertation, Mascaró (1976) develops 22 rules for Catalan phonology: he shows that 14 of these 22 rules apply only in derived contexts (he states that they are cyclic rules, but as we shall argue below, cyclic rules are restricted to applying in derived contexts only). Each of these 14 rules is neutralizing, and several of these are automatic (see pp. 17-18). Thus, (3) cannot be understood as a biconditional for there exist automatic processes that apply in derived contexts exclusively.

In the light of the preceding, one is strongly tempted to accept Kenstowicz and Kisseberth's pessimistic conclusion that 'it is not immediately clear that there is a way to predict which rules will apply only in derived contexts and which rules will apply in nonderived contexts as well. If no general principle is at work, the structural descriptions of particular rules would have to be global in nature' (p. 214). Fortunately, however, there is an interesting alternative to this. It has been argued by Mascaró (1976) that given independently motivated constraints cyclic rules will be limited to applying in derived contexts only. In order to see this we must examine what is meant by the assertion that a grammar contains rules which apply cyclically. As Mascaró remarks, at a minimum this requires that such rules obey convention (5):

(5) Given the nested constituents

$[_n \cdots [_{n-1} \cdots [_1 \cdots]_1 \cdots]_{n-1}]_n$

a set of cyclically ordered rules C will apply to the domain $[_j \cdots _j]$ only after having applied to the domain $[_{j-1} \cdots _{j-1}]$.

Convention (5), however, is not sufficiently restrictive for it allows for types of rule interactions that are much more complex than those actually observed. For example, convention (5) as stated admits a rule that applies in even-numbered constituents but not in odd-numbered constituents. Perhaps more interesting is the fact that convention (5) allows a feature switching rule to reapply on each successive constituent. Thus, it is conceivable that in a language with a feature switching rule like the vowel shift of Modern English, a word consisting of a single constituent with the stem vowel /i/ would appear on the surface as [ay]; if the same stem were part of a word consisting of two nested constituents, its vowel would be manifested as [ey]; if it were part of a word consisting of three nested constituents, its vowel would be manifested as [iy] on the surface. Moreover, the same series of outputs would be repeated in words consisting of four, five, six nested constituents, etc.[2]

Such repetitive reapplications of rules could be blocked by a restriction that would prevent rules from returning 'to earlier stages of the cycle after the derivation has moved on to larger, more inclusive domains' (Chomsky 1973:243). Chomsky, who encountered a similar problem in his work in syntax proposed the restriction (6a):

> (6) a. No [cyclic - MH] rule can apply to a domain dominated by a cyclic node A in such a way as to affect solely a proper subdomain of A dominated by a node B which is also a cyclic node.

The consequences of this restriction for phonology were explored by M.-L. Kean (1974). Although one of Kean's two examples was subsequently shown by M. Allen (1975) to be amenable to a simpler solution that did not require recourse to restriction (6a), the paper was immensely instructive in that it demonstrated in detail how general abstract restrictions on the applicability of rules such as (6a) can serve to account for rather subtle surface facts. To illustrate what is involved consider the effects of the rules (i) and (ii) on the strings (iii) in (6b) below:

(6) b. i) X Y / ___ Z ...] ii) P Q / ___ Y ...]

iii)	[...[PX]Z...]	[...[PY]Z...]	[...[P]YZ...]
i)	--	--	--
ii)	--	QY	--
i)	PYZ	--	--
ii)	blocked by (6a)	--	QYZ

Convention (6a) blocks the application of rule (6b ii) to the string [[PX]Z] on the second pass through the cyclic rules because the substring PY (from PX) is a proper subdomain of the external constituent entirely contained within the internal constituent. The same condition does not hold in the case of the substring PY in the derivation from the input [[P]YZ].

It will be readily observed that rules (6b i) and (6b ii) are structurally identical with the Finnish rules (4a) and (4b) respectively. Suppose that we assumed, therefore, that the Finnish rules are cyclic and attempted to apply them to the strings in (6c) in accordance with constraint (6a):

(6) c.	[#[vete]#]	[#[neiti]#]	[#[halut]+i#]
4a.	--	neisi	--
4b.	--	--	--
4a.	veti	--	--
4b.	blocked by (6a)	--	halus+i
Output	*veti	*neisi	halusi

As shown in (6c) we get incorrect outputs in two out of the three examples. The reason for the incorrect outputs is not the rules, but rather the constraint (6a) which determines under what conditions the rules will or will not apply. As will be shown immediately below, if the constraint on cyclic rule application is modified along lines suggested below the two rules will produce the correct output. The version of the constraint on cyclic rule application that I propose below is a combination of certain suggestions made by Kiparsky (1973:60), with others due to Mascaró (1976:9):

(7) A cyclic rule R applies properly on cycle j only if either a) or b) is satisfied:

a) R makes specific use of information, part of which is available on a prior pass through the cyclic rules, and part of which becomes first available on cycle j. There are three separate cases subsumed under a): R refers specifically to some A and B in:

i) $[_{j}XAY...[_{j-1}...B...]Z]$;

ii) $[_{j}Z[_{j-1}...B...]XAY]$;

iii) $[_{j}X[_{j-1}...A...]Y[_{j-1}...B...]Z]$

b) R makes specific use of information assigned on cycle j by a rule applying before R.

These conventions (7) generate the derivations in (8):

(8)	[#[vete]#]	[#[neiti]#]	[#[halut]+i#]
4a.	--	blocked by (7)	--
4b.	--	--	--
4a.	veti	blocked by (7)	--
4b.	vesi	blocked by (7)	halus+i
Output	vesi	neiti	halusi

Like in derivation (6c) the string [#vete#] is turned into [#veti#] by rule (4a), but unlike (6c), this string is now subject to rule (4b) by virtue of convention (7b). The rule (4b) applies to the string [#halut+i#] by virtue of convention (7a). The derivation of /neiti/ is the most interesting of all for it illustrates the important fact that given the convention (7) no cyclic rule can apply to the innermost constituent of a string. We note that (7a) requires that material from two distinct constituents be

specifically mentioned in the structural analysis of a cyclic rule R. Convention (7b) does not require this, but R must be preceded by a rule which applies by virtue of (7a) and which, therefore, cannot apply to the innermost constituent of a string. Thus, since neither of the conditions in (7) can be met by the innermost constituent of a string, no cyclic rule can apply to the innermost constituent. It follows from this that if a particular phonetic property--e.g. stress--is assigned by a cyclic rule, the rule cannot assign it to root words; instead stress on root words will have to be supplied directly in the lexical representation. It is, therefore, of considerable interest that Mascaró was able to show that in Catalan, where stress is assigned by cyclic rules, the attempts to extend these rules so that they would also supply stress to root words in their lexical representations have been failures and that stress must be supplied directly in lexical representations.[3]

Upon reading an earlier version of this paper G. N. Clements has suggested to me that convention (7) might be replaced by the following:

(9) a. A cyclic rule R applies properly on cycle j if and only if there is an immediately preceding cycle j-1 such that R must make specific use of information not present in the string at the end of that cycle.

This convention differs from (7) in that it requires cyclic rules to apply in the following two types of underlying strings which are ruled out by (7):

(9) b. i. $[_{j}Z[_{j-1}\ldots]$ XABY]

ii. $[_{j}$XA $[_{j-1}]\ldots$BZ)

I note that since application of rule R to the string AB would be ruled out if AB were part of the innermost constituent, it would seem plausible that the rule R should not be allowed to apply to the string AB when the string first appears on the jth cycle.

The only difference between the strings (9b i) and (9b ii) is that the units A and B to which the cyclic rule applies are adjacent to one another in the first, and nonadjacent in the second. I would guess that this difference between adjacent and nonadjacent constituents should not make an essential difference for a convention on rule applicability such as (7) or alternative (9). I, therefore, prefer (7) to the alternative (9). Unfortunately, neither Clements nor I have been able to think of actual examples of strings of the forms above. The difference between the formulations (7) and (9) must, therefore, remain moot at this time.

To conclude it would appear that the special conditions on rules discovered by Kiparsky are unconnected with the automatic or nonautomatic character of the rule. Their relationship to the abstractness of lexical representations is indirect: since cyclic rules cannot apply to innermost constituents they cannot be the source of phonetic properties within root morphemes, such properties will in many instances have to be directly present in the lexical representation. What is crucial about the conditions discovered by Kiparsky is that they are perfectly natural conditions on the application of the class of rules that make reference to the internal

constituent structure of words and phrases. Such rules have received somewhat less attention in phonology than they deserve, especially in recent years. Mascaró's dissertation is particularly instructive in this respect for it demonstrates in detail the important role that cyclic rules play in the segmental phonology of Catalan.

I have examined here three typical examples in which it was claimed that 'functional' rather than 'formal' considerations must be invoked in order to account for the facts, and have shown that in each case the standard 'formal' considerations provide an account that is not only adequate, but also superior to the other alternative. It would be hasty to conclude from this that 'functional' considerations are totally irrelevant in phonological research. Apart from the fact that negative results do not warrant such an inference, there is no question that 'functional' factors play a role in linguistics. If their influence is often hard to discern clearly, this is due to the fact that they are just one type among a host of factors that determine the course of linguistic evolution. By disregarding them in our research we, therefore, run the risk of missing potentially promising and enlightening lines of inquiry. The search for these factors, however, does not warrant a disregard of other--in particular, formal--factors in phonology.

NOTES

*To be published in a Festschrift for Oswald Szemerenyi, Amsterdam Studies in the Theory and History of Linguistic Science: Current Issues in Linguistic Theory (John Benjamins B.V., Amsterdam). (Preliminary Version). This work was supported in part by the National Institutes of Mental Health Grant #5P01 MH 13390-12.

[1]Kiparsky (personal communication) contests Kenstowicz and Kisseberth's analysis of Yawelmani, in particular their rule of /i/-epenthesis on which the above example crucially depends.

[2]That cyclic rules might function in this fashion was suggested by T. Bever and T. Langendoen (1963), who attempted in this manner to account for the e/o ablaut in Greek and other Indo-European languages. This ingenious idea unfortunately failed to be confirmed by the facts and has, therefore, been abandoned.

[3]The above raises questions about the SPE analysis of English stress. It should be noted that the only cases where stress needs to be supplied to the innermost constituent are the deverbal nouns of the type *condensation*.

REFERENCES

Allen, M. 1975. Vowel mutation and word stress in Welsh. Linguistic Inquiry 6. 181-201.

Bever, T. and T. Langendoen. 1963. Reciprocating cycle in the Indo-European e/o ablaut. MIT RLE Quarterly Progress Report #69. 202-3.
Chomsky, N. 1973. Conditions on transformations. A festschrift for Morris Halle, ed. by S. R. Anderson and P. Kiparsky, 232-86. New York: Holt, Rinehart, and Winston.
------ and M. Halle. 1968. The sound pattern of English. New York: Harper and Row.
Hale, K. 1973. Deep-surface canonical disparities in relation to analysis and change. Current trends in linguistics, Vol. II, ed. by T. Sebeck, 401-58. The Hague: Mouton.
Kaye, J. D. 1975. A functional explanation for rule ordering in phonology. Papers from the parasession on functionalism, 244-52. Chicago: Chicago Linguistic Society.
Kean, M.-L. 1974. The strict cycle in phonology. Linguistic Inquiry 5. 179-203.
Kenstowicz, M. To appear. Functional explanations in generative phonology. Phonology in the 1970's, ed. by D. L. Goyvaerts.
------ and C. Kisseberth. 1977. Topics in phonological theory. New York: Academic Press.
Kiparsky, P. 1973. Phonological representations. Three dimensions of linguistic theory, ed. by O. Fujimura, 3-136. Tokyo: TEC Company.
Lees, R. B. 1961. The phonology of Modern Standard Turkish. Bloomington: Indiana University Press.
Liberman, M. and A. Prince. 1977. On stress and linguistic rhythm. Linguistic Inquiry 8. 249-336.
Mascaró, J. 1976. Catalan phonology and the phonological cycle. Ph.D. dissertation, Department of Linguistics, MIT. Distributed by Indiana University Linguistics Club, Bloomington.
Zimmer, K. 1975. Some thoughts on likely phonologies for non-ideal speakers. Papers from the parasession on functionalism, 556-67. Chicago: Chicago Linguistic Society.

FUNCTIONALISM AND FUNCTIONAL EXPLANATIONS IN PHONOLOGY

Jonathan Derek Kaye
Université du Québec à Montréal

0. In this paper I will do my best to shed a little light on the rather murky area of functionalism as it is applied in phonology. In part one I will briefly sketch two of the major currents in the functional approach. In addition I will attempt to show what relationship each of these functionalist approaches bears to current phonological practice.

The first variety of functionalism to be considered (all too briefly) is what I call 'natural functionalism' in the sense of Donegan and Stampe (1978).[1] The second type of functionalism is, I will argue, a mixture of two rather different elements: questions of substantive conditions in phonology and a posteriori functional explanations. This latter sort of functionalism is associated with Kiparsky's phonology ca. 1970. It is also to be found as applied to syntax in Chomsky and Lasnik (1977) and Chomsky (1977). For identification purposes and for lack of a better label I will call this 'a posteriori functionalism' recalling its essentially parasitic status with respect to normal linguistic hypothesis formation.

These approaches will be discussed from the point of view of generative phonology. Naturally my conclusions should be understood in this context. Other theoretical frameworks may lead to quite different results. Because of the above-mentioned orientation I will devote more attention to the question of a posteriori functionalism.

Part two of this paper will illustrate the way a posteriori functional explanations figure in phonological analyses. A problem from Algonquin (Kaye 1978a, b; Kaye and Nykiel 1978) is discussed and a substantive condition on evaluating phonologies is proposed. A functional explanation is then suggested to account for the substantive condition.

1. One of the major problems in discussing functionalism is that there are just about as many different functionalisms as there are linguists who use the term.[2] I do not wish to embark on an exhaustive historical study of the origins, varieties and evolution of functionalism and so I will sketch what I hope is a reasonably accurate, albeit of necessity much abbreviated description of natural functionalism. Since definitions of this approach are exceedingly difficult to find I will present my own. Natural functionalism is characterized by the assumption that there is a direct and clear link between what is judged to be an essential feature of the object of study (such as its function) and theories about that object. Expressed somewhat differently it is claimed that starting from this essential element we may deduce aspects of the theory which is to describe it. This position may be summed up by the expression, 'form follows function'.[3]

The functionalism of André Martinet seems to me to be a good example of this position. Martinet (1972:7) states '. . . le langage humain est, au fond, fait pour la communication, et qu'il se façonne du fait qu'il sert à la communication.' For Martinet the essential feature or function of language is communicative. This notion can be seen in his discussion of push chains. Potential mergers are avoided by successive phonetic shifts, each shift creating another potential merger and provoking a further phonetic shift. This state of affairs follows from the idea that the function of language is communicative and the fact that a merger would run counter to this function to a greater or lesser extent. A phonological opposition once lost is no longer available to a language to express semantic distinctions and communication is accordingly rendered more difficult.

A similar situation is found in functional syntax. When English lost case endings as a means to express grammatical relations, word order came to fulfill this role. It was reasoned that grammatical relations need to be expressed in order for communication to take place. If one means of accomplishing this (inflexional endings) ceased to be available for one reason or another, it was necessary that another mechanism be brought in (word order).

The theory of 'natural phonology' as proposed by Donegan and Stampe, while differing enormously from Martinet's functionalism in both the nature of its claims and its internal organization, seems to me to share with it a certain philosophical base. In describing their theory Donegan and Stampe (1978:1) state that, 'its basic thesis is that the living sound patterns of languages, in their development in each individual as well as in their evolution over the centuries, are governed by forces implicit in human vocalization and perception.' The essential feature of language mentioned in the characterization of natural functionalism given above is not its function for Donegan and Stampe. Strictly speaking their theory is not functional in the above sense. The term 'essential feature' rather than 'function' was intended to be broad enough to include natural phonology within the domain of natural functionalism. Precisely what this essential feature is, is made clear in the following passage: 'The natural subject matter of an explanatory theory includes all and only what the theory can, in principle, explain. In the case of natural phonology this means everything that language owes to the fact that it is spoken' (1978:4). While communication is the key element in Martinet's functionalism, Donegan and Stampe focus on the oral nature of language.[4] The relationship between these key concepts and the theories associated with them appears to be quite similar. As the communicative function of language is supposed to be evident in the theory of Martinet so the oral nature of phonology should be evident in Donegan and Stampe's natural phonology.[5] In their discussion of the question of innateness they give a further indication of the functional nature of their theory: 'The issue of innateness, despite all the debate it has aroused is entirely besides the point. What we want to know, whether it is innate or whether it is universally acquired, is <u>why</u> . . .' (1978:5, original emphasis).

The natural functionalist seeks to construct a theory which provides its key concept as the answer to all the questions of the sort raised in the above citation, i.e., natural phonology is expected to explain the

universal presence of a given trait (innate or not) because of the oral nature of language; Martinet would presumably respond to the question with a theory offering language's communicative functional as at least part of the reply.

Certain anthropological and sociolinguistic approaches appeal to extralinguistic features as being essential to the formulation of adequate linguistic descriptions. Thus, it is often claimed that language cannot be properly analyzed outside of its cultural context, or else that language must be considered as a social instrument and theories which purport to explain it cannot ignore this fact. The details of language's social function or its cultural context are not always clear, nor is their exact status in the formulation of linguistic hypotheses evident. Such notions appear to play a similar if not identical role to those mentioned in the discussion of Martinet functionalism and natural phonology.

A number of questions are raised by the sort of functionalism sketched above. First, the metatheoretical status of natural functionalism and its accompanying essential features is open to at least two interpretations. One may view the problem of functionalism as essentially methodological. Natural functional considerations of the type discussed above may serve as heuristic aids in formulating linguistic hypotheses. Under this interpretation the functionalism of Martinet, to take one example, would instruct us that language's communicative function is a useful concept to keep in mind when formulating linguistic hypotheses. The second interpretation is to consider the function (essential feature) of language to be the object of study. A theory is evaluated on the basis of the extent to which it explains or reflects the essential feature that is supposed to characterize language. Thus, a theory explaining why or demonstrating that language is a social, cultural, or communicative instrument would be considered as more highly valued than one which treats such notions as irrelevant. A phonological theory which does not reflect the essentially phonetic nature of language would be regarded as inadequate. I will argue that neither interpretation of natural functionalism constitutes a viable alternative to standard generative phonological practice. There are two general problems associated with both of these interpretations. First, it is not at all obvious what function a functional theory of language is to be based on. Indeed, one encounters almost as many brands of functionalism as there are writers on the subject. Furthermore, if the function of a language is all that crucial either for methodological reasons or as an object of study, it follows that functional theories should differ wildly according to what has been chosen as language's function. But then how are we to evaluate these competing theories? Surely this must be done in exactly the same way that any nonfunctional theory is to be evaluated.

It must be remembered that even the most intuitively 'obvious' notion of language function is not uncontroversial. Thus, Martinet's position that the function of language is communication is contested by Chomsky in several recent works. Although not taking a strong position on the question, Chomsky has suggested that the expression of thought is at least as strong a candidate as communication. It is not my purpose here to try and settle this issue, rather I wish to point out that the function

of language is still a very open question and one is by no means constrained to accept any given definition in the absence of supporting evidence. Up to now it is just such evidence that has been lacking.

The second general objection to natural functionalism is more serious. Even if somehow we could all agree on what the function of language really is, it is not at all clear that this would change anything about the way we would do linguistics. There is no reason to suppose that there is any clear or direct line between form and function in language. If this is so, then using language function as a starting point would appear to be methodologically unsound. Evaluating theories on the basis of their shedding light on an a priori defined linguistic function (e.g. rejecting the EST because it is unrevealing of language's social function) seems misguided.

It is instructive to study the formulation of theories in other disciplines to emphasize this point. One can immediately exclude many of the 'hard' sciences from discussions of functionalism as defined above. It would be absurd to think of a priori functional explanations for astronomical, physical, or chemical phenomena. These sciences have survived without requiring natural functionalism either to guide them to fruitful hypotheses or to form the framework for eventual explanations of the phenomena within their domain. It is only when we approach the biological sciences that a discussion of functionalism makes any sense at all. Here the discovery of the structure of DNA is quite illuminating. In reading Watson's account of how he and Crick came up with their proposed structure (Watson 1968), it is evident that not only did the function of DNA play no role in its analysis but that one of Watson and Crick's preoccupations was that its structure would not be 'pretty': 'pretty' in the sense of revealing something about how DNA worked. Thus, Watson writes (1968:188), "There had been far too many days when Francis (Crick/JDK) and I worried that the DNA structure might turn out to be superficially very dull, suggesting nothing about either its replication or its function in controlling cell biochemistry" (emphasis mine/JDK). To the delight of Watson and Crick, DNA's structure turned out to be quite pretty indeed. What I wish to stress here is that there was a chance that DNA's structure could have been 'dull' in Watson's sense of the word. This case is interesting because there was general agreement about DNA's function before its structure had been discovered and this knowledge was not particularly helpful in its analysis. The application of this point to linguistics is clear: there is no reason to suppose in advance that linguistic structure bears any direct relationship to any function, real or imagined. I will now turn to specific criticisms of the two interpretations of natural functionalism.

Natural functionalism may be interpreted as requiring or suggesting that a language's function be taken into account in formulating hypotheses about its structure. From an epistemological point of view the above statement is to be rejected. In evaluating hypotheses the method that was used to arrive at them is irrelevant. Hypotheses are evaluated as to how well they explain the phenomena found within their domain and how they compare to other hypotheses with similar domains. How one came by these hypotheses does not figure in the evaluation.[6]

We could also consider this methodological interpretation as a heuristic device: while not being absolutely necessary to the formulation of linguistic theories, it may be a good idea or a useful suggestion to follow such a procedure. I have not seen any evidence that such is the case. The theory of generative phonology has progressed normally over the years without any particular emphasis on functional considerations,[7] and no clear examples come to mind where functional considerations have helped to overcome a particular problem in this framework. I conclude that the case for natural functionalism as a methodological tool remains unproved, indeed without any apparent support of any kind. We may now turn our attention to the claim that linguistic theory must explain or incorporate the particular function or essential feature assigned to language by the brand of functionalism in question. Natural phonology requires that linguistic theory include 'everything that language owes to the fact that it is spoken.' It is perfectly appropriate that one should choose this as the object of study of phonology. I fail to see, however, that one should accord this approach some prior claim to legitimacy. It seems to me that natural phonology is a reflection of the research that happens to interest its adherents. Failure to share one's research interests is hardly a crushing criticism. In as much as natural phonology and generative phonology share common objectives, one may compare them. If they seek to achieve fundamentally different goals, any comparison or criticism is pointless. Language can certainly be described from a social, cultural, or communicative point of view. It does not follow that such descriptions or theories have anything to do with a linguistic theory that seeks to describe whatever internalized knowledge characterizes the set of speakers of some language and what universal principles underlie the set of all languages. The functional approach described by Heath (1978:89) attempts to ". . . describe a set of functions which, by means of different combinations of formal units and with inevitable variations in (sociocultural) environmental detail, play fundamental roles in shaping the formal grammars of individual languages.' There is no more reason to accept this view of linguistic theory than any other. Ultimately the results emanating from such programs and further lines of research suggested by them will determine their success. Again, I can see no reason to accord any prior claim to research programs incorporating functionalist orientations. To conclude this extremely brief discussion of natural functionalism, I see no role for such an approach in the development of phonological theory. It appears epistemologically unsound and offers no clear methodological advantages. A priori theoretical constraints of this sort seem unduly confining or else too vague to be of much use. As for their role in defining research areas, natural functionalism has not proved any more fruitful than nonfunctional approaches. Finally, functionalism has not played any obvious role in the development of other (nonsocial) sciences.

A rather different functional approach has come into prominence in the last ten years. P. Kiparsky has been the dominant figure in this approach with C. Kisseberth, M. Kenstowicz, Y.-Ch. Morin, K. Hale, and many others normally associated with generative phonology also making contributions to this area.[8] I will show that this current is radically different from natural functionalism and for the most part is really a confusion of several distinct questions.[9]

Generative phonology attempts to explain, among other things, how a child exposed to a limited set of primary linguistic data arrives at an extremely rich phonological system in a remarkably short period of time. It has been assumed that the child had an enormous 'head start' in the acquisition process: a way of selecting a phonological grammar from the set of all those that are compatible with the primary linguistic data to which he/she was exposed. This 'head start' was characterized in generative phonology by the evaluation metric: a diverse set of properties that were reflected by the theory's notational system such that more general or natural rules and systems were less costly (required fewer symbols to express) in this system. Towards the end of the sixties Kiparsky came up with several properties that were not conveniently reflected in the notational system but that also seemed to characterize more natural linguistic systems. It was claimed that such properties were also used by children acquiring a language in ways analogous to the evaluation metric. Kiparsky (1971) termed these conditions substantive as opposed to formal. These substantive conditions which included such notions as opacity, abstractness, and paradigm regularity were no more functional in nature than natural classes, feature hierarchies, and markedness conventions. The latter could be read off directly from the notational system while the former could not. Nothing in principle would prevent the eventual refinement of the notational system so that say, opacity could likewise be directly represented. At that point opacity would be a formal rather than a substantive condition on phonologies but it would in no way change any functional characteristics it might have. It just so happens that functional explanations became associated with these substantive conditions although there is no logical link between them. To take one example, Kiparsky showed that rules became less highly valued to the extent to which they had become opaque. Kiparsky's definition of rule opacity naturally accompanied this claim. At this point there is nothing whatever that would lead one to consider Kiparsky's contribution as functional. Kiparsky went on to speculate as to why language should favor transparent rules over opaque ones and suggested that transparent rules are easier to learn. It is at this point that a functional element entered the discussion. Since functional notions of this sort are made only after the postulation of substantive conditions on phonology, I have called this sort of functionalism, 'a posteriori functionalism'. In fact, as I will try to show below, this is not a distinct theoretical orientation from normal generative phonology.

To begin with, two points must be made clear: the conjectural nature of the functional claims and the fact that the validity of the hypotheses on which they are based is completely independent of such functional statements. Functional statements along the lines of Kiparsky's assertion that transparent rules are easier to learn, are generally offered as a sort of 'coda' that follows normal phonological arguments justifying the substantive conditions in question (in this case rule opacity). There is an (implicit) appeal to common sense but no effort is devoted to justifying these sorts of claims. Notice also that the validity of the functional statement has no bearing on the validity of the theoretical status of the substantive condition on which it was based. Suppose, for example, that we could somehow prove that opaque rules were not more difficult to learn than transparent ones. This would in no way imperil the claim that

grammars with opaque rules are less highly valued than those without them. This latter claim was based on evidence completely independent of any functional considerations.[10] The data for which rule opacity was set up to account would remain untouched by the success or failure of any functional explanation. The lack of any functional motivation for a substantive condition would simply indicate that this particular aspect of linguistic structure was not 'pretty' in the sense of Watson discussed above.

Susumu Kuno was one of the early practitioners of a posteriori functionalism. Kuno (1972) suggested that many of Greenberg's word order universals follow if language processing involved a push-down storage device at some point. Again, notice that this claim was made only after Greenberg had postulated his universals. Greenberg's observations about word order were not made with functional considerations in mind, nor does their validity rest on the success of a posteriori functional explanations.

Chomsky and Lasnik make repeated reference to functional explanations of this kind. For example they state (1977:470), 'As we have observed, "functional" considerations are quite appropriate in the study of language evolution. We have pointed out that these filters have at least one striking effect: they sharply reduce the range of possible outcomes from well-formed base-generated structures, while still permitting at least one in each case. Thus they narrowly restrict the association of deep and surface structures.' Chomsky and Lasnik certainly did not set out to devise a linguistic system with the above-mentioned property (i.e. reducing the range of possible outcomes from well-formed base-generated structures). This result was the result of normal syntactic research. It just so happened that the resulting model was pretty enough to allow these functional properties to be read off of it. Even in this case, however, the explanation is not quite functional. One could propose a general condition on grammars to the effect that they are more highly valued if they possess the property of having reduced ranges of outcomes from well-formed structures. Predictions involving linguistic change, interference phenomena and perhaps some absolute limits on the range of possible grammars could follow from such a condition on grammar. It must be noted that there is no a priori reason to believe that such a condition exists. It could only be justified in the same way that the opacity condition, the subjacency condition for movement rules or the filters themselves were justified. Given the existence of this substantive constraint one might then offer a functional explanation (it makes processing easier, etc.) to account for its presence. In any event the position of those functional explanations in linguistic theory is quite tenuous. 'The functional explanation . . . holds, if at all, at the level of evolution of the species. It does not relate to explanatory adequacy . . .' (Chomsky and Lasnik 1977:437).

Returning to phonology, Halle in his contribution to this volume has suggested that all functional considerations should reside in the simplicity metric. In other words, the destinction between formal conditions on phonology (say, rules involving natural classes are more highly valued than those without such classes) and substantive conditions (say, rule

opacity) should be eliminated. This could be interpreted to mean that all conditions ranking phonologies should be formulatable in the notational system, either in the traditional way, higher ranking phonologies being less 'costly' in terms of the number of symbols needed to express them, or by extending the formalism and adding ancillary conditions.

The former conditions include the natural classes, abbreviatory devices and so on. The system was subsequently enriched to include markedness, rule variables, metric structures, and so on. There appears to be no reason to distinguish these two classes and in that sense I agree with Halle's suggestion. Furthermore, it seems a reasonable proposal to attempt to incorporate substantive conditions such as rule opacity and recoverability into the notational system. I see no problem in such an endeavor and at worst it would have as a consequence the elimination of an arbitrary distinction in current generative phonology. Functional explanations may be just as easily proposed for, say, markedness, which is currently considered a formal condition, as for rule opacity, which is a substantive one. Thus Halle's proposal would leave unchanged the functional status of generative phonology. It would have the beneficial effect of disentangling functionalism from substantive conditions on phonology. In the following section I will illustrate the sort of a posteriori functionalism discussed in the previous paragraphs.

2. In this section I will discuss a substantive condition on phonology and show its role in the analysis of a particular phonological problem. I will indicate how such questions relate to the a posteriori functionalism discussed in the preceding section.

In some recent articles (Kaye 1974; 1975) I discussed the notion of recoverability in phonology. Very roughly, recoverability is a limit on the amount of surface ambiguity of a phonological form. What this means is that given a phonetic form and the rules, the inventory of phonemes and (deep) phonotactic constraints of the language in question, the number of possible sources of the phonetic form may be computed. For example, in German given a phonetic string ending in [-t], given the rule of final devoicing and given the fact that morphemes may end in *-t* or *-d*, the string has two possible sources, viz. /-t/ or /-d/. Only comparison with morphologically related forms containing suffixes will indicate which source is to be chosen. In the absence of such related forms one would choose /-t/ as the underlying final consonant since German final devoicing is a neutralization rule. Recoverability then evaluates phonologies as more highly valued if such ambiguity falls within certain limits. In the example below I will illustrate the application of this principle to (deep) phonotactic constraints. Lac Simon Algonquin contains both a word structure constraint (WSC) which is rendered completely opaque by a phonological rule.[11]

Lac Simon Algonquin has the WSC given in (1) below.

(1) Algonquin WSC: No word may begin with a cluster or a voiceless consonant.

In addition the following data (2) indicate that Algonquin has acquired a

rule of initial devoicing (3).

(2)

pi:goška: 'it breaks'	*ki:bi:goška:* 'it broke'
ka:zo:tə́m 'he hides'	*nə́ga:zo:ta:n* 'I hide'
te:sə́biwa:gin 'chair'	*kə́de:sə́biwa:gə́n* 'your chair'
sə́gə́swa: 'he smokes'	*ta:zə́gə́swa:* 'he would smoke'
či:ma:n 'canoe'	*oǰi:ma:n* 'his canoe'
šo:škose: 'it slides'	*nə́žo:škose:* 'I slide'

(3) Initial Devoicing C → [-voice] /## ___

While not altogether a complete justification for the analysis (which is irrelevant in any event for the point under discussion), the following data indicate that we are dealing with a rule of initial devoicing rather than medial voicing.

(4)

a.	*kə́t̲ga:n* 'cultivated field'	*kə́čə́* 'much, big, great'
	a:sə́bi: 'net'	*pi:sa:də́k* 'in height'
	a:nə́pi:c 'when?'	*nə́do:žə́to:n* 'I make it'
b.	*nə́-kə́n* 'my bone'	*o-se:z-ə́n* 'his brother'
	kə́-šogona: 'your nose'	*nə́-ka:t* 'my leg'
c.	*a:sə́bi:* 'net'	*a:sə́bi:-k̲a:k* 'in the net'
	pa:pa:se: 'woodpecker'	*pa:pa:se:-š̲ə́š* 'little woodpecker'

(4a) shows that both voiced and voiceless obstruents occur word-medially although only voiceless obstruents occur word-initially at the phonetic level. (4b) shows that dependent noun stems may also begin with voiceless obstruents. These forms are preceded obligatorily by prefixes and thus cannot violate (1). Finally, (4c) shows that suffixes may also have initial voiceless obstruents.

The status of nonalternating (i.e. always initial) morphemes is crucial for this analysis. If one assumes that rule (3) is neutralizing, nonalternating obstruent-initial morphemes as in (5) would be required to have initial voiceless obstruents in their underlying representations. If, on the other hand, rule (3) is assumed to be allophonic, even nonalternating obstruent-initial morphemes would have underlying initial voiced obstruents.

(5) Obligatorily initial morphemes with initial obstruents

kə́- '2nd person prefix'	*paiə́me:* 'it is necessary'	*ka:-* 'relative marker'
ki:špə́n 'if'	*te:də́go:* 'much'	*pi:č* 'and then'

This situation is somewhat reminiscent of the much discussed German case. The two languages are compared in (6) below.

(6) The status of rule IV

LAC SIMON			GERMAN	
INITIAL	NON-INITIAL		FINAL	NON-FINAL
voiceless	voiced /voiced/	/voiced/	voiceless	voiced
po:de:	*ki:bo:de:*		*rat*	*rɐdɔr*
'he makes a fire	he made a fire'		'wheel'	'wheels'

voiceless	voiceless	/voiceless/	/voiceless/	voiceless	voiceless
NO FORMS EXIST				*rat* 'counsel'	*rɛtə* 'counsels'
voiceless	always initial	/???/	/voiceless/	voiceless	always final
pa:ńme: 'it is necessary...'				*unt* 'and'	

Both German and Algonquin have voiced-voiceless alternations and both have forms which always occur in the suitable context for the devoicing rule to apply (i.e. initially in Algonquin and finally in German). Algonquin differs from German in that there are no examples where a voiceless obstruent occurring in initial position is also voiceless when the same morpheme follows a prefix. This is, or course, simply a reflection of (1). The question now centers around the forms of (5). In (7) the two possible solutions are sketched. Solution I assumes that the forms of (5) have underlying voiced obstruents initially while solution II assumes they are voiceless. Since both solutions are compatible with the data we must look elsewhere for evidence in support of one or the other analysis. Furthermore, we must find a principled basis for preferring or requiring one solution over the other.

(7) Solution I

/gi:špin/ /de:digo:/ /bi:č/

Rule IV: [+cons, -son] → [-voice] /## ___

Word Structure Constraint: No word may begin with a voiceless consonant.

Solution II

/ki:špin/ /te:dɨ́go:/ /pi:č/

Rule IV: [+cons, -son] → [-voice] /## ___

The principle difference between the two solutions is the presence or absence of the WSC. Evidence from the nativization of loan words favors the former solution.

(8) Loan Word Behaviour

a. *pa:na:n* 'banana' (<banana) *nɨ́ba:na:nɨ́m 'my banana'* /ba:na:n/
pa:štone: 'American' (<Boston) *ka:ba:štone:na:gonɨ́č* 'he who looks like an American' /ba:štone:/

b. *pɨ́tɨ́k* 'potato' (<petac) *nɨ́bɨ́tɨ́kɨ́m* 'my potato' /bɨ́tɨ́k/
toma:do:s 'tomato' (<tomatoes) *nɨ́doma:do:zɨ́m* 'my tomato' /doma:do:z/
ko:fi:ke: 'he makes coffee' (<coffee) *nɨ́go:fi:ke:* 'I make coffee'/ go:fi:ke:/

In (8a) loan words with an initial voiced obstruent in their source language are presented. When preceded by a prefix the initial obstruent retains its voiced status. In absolute initial position it is realized as voiceless. These forms have no bearing on the status of the WSC (1) but rather show that rule (3) is still operative. The forms of (8b) begin with voiceless consonants in the source language. In Algonquin they are voiceless in initial position. In noninitial position, however, they are voiced. Since we have already stated that there is no rule of medial voicing, we must assume that these forms were interpreted as having an initial voiced obstruent. The only apparent reason for such an interpretation is the

existence of the WSC (1), ruling out word-initial voiceless obstruents as possible underlying forms.

Evidence from loan word phonology favors solution I over solution II. What feature of this solution renders it more highly valued than the other solution? Both solutions require the presence of rule (3). The only difference in the formal structure of these solutions is the presence of a phonotactic constraint in solution I. By incorporating a substantive condition into our phonological theory we can require the selection of solution I over solution II. The substantive condition in question is given below:

(9) All things being equal, phonologies with phonotactic constraints on possible underlying forms are more highly valued than those without such constraints.

If this turns out to be a correct observation about ranking phonologies one may well turn to the question of why phonological theory has a condition like (9): why should analyses with phonotactic constraints be selected in the acquisition process over those without such constraints, even in the face of apparently contradictory surface forms? In other words, is there a functional explanation for (9)? One striking property of this condition is to limit the number of underlying forms and consequently to facilitate the task of associating a given phonetic form with its lexical representation. Indeed we may view (9) as simply a special case of a more general condition of recoverability. Note that the case for phonotactic constraints as well as other manifestations of recoverability may simply be another aspect of Chomsky and Lasnik's functional explanation mentioned above. Filters had the property of reducing the number of outcomes associated with a given base-generated structure. Phonotactic constraints have the property of reducing the possible underlying sources associated with a given phonetic string. It would be most tempting to look for a general functional explanation which would incorporate both cases.

The purpose of the above illustration was to show the role of (a posteriori) functional explanations in the current practice of generative phonology. It should be noted that things go on pretty much as they did in the past with functional considerations playing no role in the quest for explanatory adequacy. Like Watson and Crick, phonologists will go on with their analyses with the hope that what results will be pretty.

NOTES

[1] 'This is a natural theory in the sense established by Plato in the Cratylus, in that it presents language (specifically the phonological aspect of language) as a natural reflection of the needs, capacities, and world of its users, rather than as a merely conventional institution' (1978:2; original emphasis).

[2]See Baron (1974) for discussion of the problem.

[3]Heath (1978:92). '(Although) in this framework we take function as prior to form . . .'

[4]It must be stressed that Donegan and Stampe are only talking about the phonological aspect of language. Martinet is dealing with all aspects of language.

[5]If my understanding of natural phonology is correct, anything remotely resembling a phonological component in a sign language such as ASL should be pure coincidence.

[6]See Feyerabend (1978) for a discussion of this question.

[7]Cf. the discussion of a posteriori functionalism below.

[8]See also Morris Halle's contribution to this volume.

[9]I am one of the parties guilty of this confusion. See Kaye (1975).

[10]Kiparsky's arguments for this rule opacity involve evidence from sound change, specifically rule reordering.

[11]This analysis has proved to be rather controversial. The reader is referred to Kaye (1978a; 1978b) and Kaye and Nykiel (1978) for the argumentation.

REFERENCES

Baron, N. S. 1974. Functional motivations for age grading in linguistic innovation. Historical linguistics I, ed. by J. Anderson and C. Jones, 33-64. Amsterdam: North Holland.

Chomsky, N. 1977. Dialogues avec Mitsou Ronat. Paris: Flammarion.

------ and H. Lasnik. 1977. Filters and control. Linguistic Inquiry 8.425-504.

Feyerabend, P. 1978. Against method. London: Verso.

Heath, J. 1978. Functional universals. Berkeley Linguistic Society 4.86-95.

Kaye, J. 1974. Opacity and recoverability in phonology. Canadian Journal of Linguistics 19.134-149.

------. 1975. A functional explanation for rule ordering in phonology. Papers from the parasession on functionalism, 244:52. Chicago: Chicago Linguistic Society.

------. 1978a. On the alleged correlation of markedness and rule function. Current phonological theories, ed. by D. Dinnsen. Bloomington: Indiana University Press.

------. 1978b. Abstractness recoverability and phonotactic constraints. To appear in Phonology in the 1970's, ed. by D. Goyvaerts.

------ and B. Nykiel. 1978. Loan words and abstract phonotactic constraints. Paper read at the 16th International Polish-English Contrastive Linguistics Conference, Boszkowo, Poland.

Kiparsky, P. 1971. Historical linguistics. A survey of linguistic science, ed. by W. Dingwall, 576-642. College Park: University of Maryland.

Kuno, S. 1972. Natural explanations for some syntactic universals. Mathematical linguistics and automatic translation. Report no. NSF-28 to the National Science Foundation. Harvard University.

Watson, J. 1968. The double helix. New York: Atheneum.

INTENTIONALITY AND THE USE OF LANGUAGE

John R. Searle
University of California at Berkeley

It is possible to make a reasonably clear distinction between intentionality with an s and intentionality with a t. Without putting too fine a point on it one can say that intensionality-with-an-s is a property of a certain class of sentences. A sentence is intensional if literal utterances of it have at least one interpretation where they fail to satisfy one or more of the standard tests for extensionality. The two tests most relevant to the present discussion are these: if existential generalization over the occurrence of referring expressions is not a valid form of inference or if the sentence fails to allow the substitution salva veritate of expressions which normally have the same reference, then it is intensional-with-an-s. Thus for example, the sentence 'John is looking for the lost city of Atlantis' has at least one literal use to make a statement which does not entail that there is a lost city such that John is looking for it. And the sentence 'The sheriff believes that Mr. Howard is an honest man' has a literal use to make a statement which together with the true statement that Mr. Howard is identical with Jesse James does not entail that the sheriff believes that Jesse James is an honest man. Because intensional sentences normally derive their intensionality from the occurrence of certain expressions it is possible to speak not only of intensional sentences, but also of intensional verbs, intensional contexts, etc. It is also possible to speak of intensional statements and intensional propositions. No doubt these two tests need some refinement to enable them to cope with sentences other than those used to make statements but the basic idea of intensionality-with-an-s seems reasonably straightforward.

When it comes to intentionality-with-a-t it is much harder to say exactly what it is we are talking about. It is common to say that intentionality-with-a-t is not primarily a property of sentences, but a property of some or perhaps even all mental phenomena. It is the property which mental states have of being directed at objects or states of affairs. Thus for example, a belief is always a belief <u>that</u> such and such is the case, a fear is always--or at least in general--a fear <u>of</u> something, a desire is always a desire <u>to have</u> something or that something be the case. Just so we have some fairly clear idea of what it is we are talking about, I propose the following as a rough preliminary test for intentionality-with-a-t. A mental state is an intentional state if and only if the specification of the content of that mental state requires the specification of some object or state of affairs which is not identical with that mental state. On this test pains and aches and at least some cases of anxiety are not intentional, whereas beliefs, hopes, expectations, and desires are. A specification of the content of my pain is just a further description of the pain; but the specification of the content of my belief, hope, expectation, or

desire must specify what it is that I believe, what I hope for, what I expect, and what I desire. All that a pain has to have to be a pain is to have a certain feel to it, certain phenomenal qualities, but for belief, hope, expectation, and desire, something more is required: the questions 'what do you believe, what do you hope for, what do you expect, what do you desire?' must have answers, if the agent can be said to have a belief, hope, desire, or expectation at all; and those answers will specify objects and states of affairs that are not identical with the mental states.

In Wittgenstein's jargon, pains and aches have causes but not targets, whereas love and hate, belief and desire have both causes and targets, and in each case the cause may or may not be identical with the target. I will call such mental states intentional states, and the objects and states of affairs at which they are directed intentional objects. The notion of an intentional object is a frequent source of confusion in philosophy and I will have more to say about it later on. We will see that for the so-called propositional attitudes it is not really necessary at all. For the present it is important to note that on this criterion not all mental states are intentional states; only those which require the specification of an intentional object are intentional states. It is also important to note that this criterion is not intended as an analysis of the notion of intentionality. If so, it would be hopelessly inadequate since it rests on several unexplained and obscure notions: in what sense does a mental state have a 'content' and what is meant by saying the 'specification' of that content 'requires' the specification of something else?

This test for intentionality-with-a-t is designed to isolate those mental states which are in some sense directed at objects and states of affairs from those which are not; but even assuming that it is successful in doing that, it raises but so far leaves unanswered the hard questions about intentionality. The two most pressing questions raised by the criterion are these: the statements specifying the object or state of affairs in the specifications of the intentional state are not like ordinary statements containing specifications of objects and states of affairs because they do not require the existence of the objects or state of affairs in order to be true. It can, for example, be true that John expected it would rain and that he worships God without it being true either that it rained or that God exists. Indeed, the sentences which specify intentional states are themselves intensional-with-an-s on both our criteria for intensionality-with-an-s. So our first question is: why is that so? How does intentionality-with-a-t give rise to intensionality-with-an-s? The second question is this: what exactly is the relation (or set of relations) between intentional states and their objects, and what is it about this relation that requires the specification of an intentional object in the specification of an intentional state? I think an answer to the second question will enable us to answer the first and to clarify the notion of an intentional object.

In discussing the second question, Wittgenstein frequently makes remarks like the following 'If someone could see the expectation itself--he would have to see what is being expected' and 'The representation of a wish is eo ipso the representation of its fulfillment' (cf. Wittgenstein 1967a, 1967b). Wittgenstein himself and Kenny in his book on Wittgenstein

claim that there is an 'internal relation' between the intentional state and its intentional object. But this characterization is not much help to us unless we are told exactly what is meant by the notion of an internal relation. If for example, someone tells us that there is an internal relation between being a triangle and being three-sided, it is possible to make some fairly clear sense of the notion of an internal relation: a logically necessary condition of being a triangle is being three-sided. But in that sense there clearly is no internal relation between my belief that it will rain on Wednesday and the state of affairs of its raining on Wednesday. The proposition that it will rain on Wednesday neither entails nor is entailed by the proposition that I believe that it will rain on Wednesday. It seems that the only sense we can give to the notion of an internal relation in this case is one we have already spelled out in our test for intentional states--the specification of an intentional mental state requires a specification of its intentional object. But that still leaves us with our second question unanswered. What is it exactly about intentional states that relates them to their objects? Mackie (1975:48) poses our second question as sharply as anyone: 'At last we are beginning to get into focus the real puzzle about intentionality,' he writes, 'Is it not strange that there should be one state of affairs (Tom's state of believing, hoping, fearing, or whatever it may be) that requires for its adequate description the partial, incomplete, selective, indeterminate, description of a quite different and so far merely possible state of affairs?' I think there is a fairly simple answer to Mackie's question. To see this answer imagine a parallel question about statements and speech acts generally. Suppose someone said, 'Is it not strange that there should be one state of affairs (Tom's statement that it is raining or his order to Bill to leave the room) that requires for its adequate description the partial, incomplete, etc., description of a quite different and so far merely possible state of affairs (the state of affairs that it is raining or that Bill leaves the room).' But there is nothing puzzling about this requirement once you recognize that a statement is precisely a representation of a state of affairs that is asserted to exist and an order is a representation of a state of affairs which the hearer is ordered to bring into existence. Any speech act with a propositional content contains a representation of some object or state of affairs; and in the same way, I wish to argue, all intentional states are representations of objects and states of affairs. The reason that the specification of my belief that it is raining requires a specification of the state of affairs that it is raining is exactly the same reason that the specification of my statement that it is raining requires a specification of the state of affairs that it is raining: in each case the one is a representation of the other.

The answer to our second question then, how does the intentional state relate to the intentional object and what is it about the relation that requires that a specification of one involves a specification of the other, is simply that the intentional state contains a representation of the intentional object in the same sense that speech acts contain representations of objects and states of affairs.

Because an intentional state contains a representation we can now give a clear sense to the notion that it is 'internally related' to the object it represents: Any representation is internally related to its object in

the sense that it could not be that representation if it did not have that object. Thus for example, an identity criterion of my belief that it is raining is that it must have as its intentional object the state of affairs that it is raining.

I don't much like having to invoke the notion of 'representation' here because it has so little explanatory power, but it is at least correct and it does enable us to answer both our second and now our first question, How does intentionality-with-a-t give rise to intensionality-with-an-s? Sentences about intentional states are at least in part about representations. That being the case, their truth conditions will sometimes depend on features of the representation and not entirely on features of--or even on the existence of--the object represented. To be looking for the lost city in Atlantis is to have a representation that one seeks to instantiate, and the statement that one has such a representation can be true even where there is no such instantiation. And similarly to believe that Mr. Howard is an honest man is in part to have a mental representation associated with the name 'Mr. Howard' which may be quite different from the mental representation associated with the name 'Jesse James', hence the failure of the substitutability of such names even where Mr. Howard is identical with Jesse James.

This conception of intentionality as representation will also enable us to get clear about the nature of intentional objects. People often talk as if intentional objects had some peculiar ontological status and had to be distinguished from actual objects. But the intentional object of a mental state is just the actual object or state of affairs represented by an intentional state. If there is no such object or state of affairs then the intentional state does not have an intentional object though it does still contain a representation. We need therefore to distinguish the representative content of a mental state from the intentional object of that mental state. If John loves Sally and believes that it is raining then the intentional object of his love is Sally, the actual flesh and blood Sally and not some mental phenomenon, and the intentional object of his belief is the state of affairs in the world that it is raining and not the proposition that it is raining. In order that his love should be of Sally and his belief be that it is raining he must have some representation of Sally and of the state of affairs that it is raining. But the object of his intentional states is not these representations, rather the intentional states are directed at their objects by way of their representative content. A proposition one might say is not the object of a belief, it is the content of the belief. The oscillation between the extensional and the intensional reading of statements about intentional states is precisely an oscillation in the extent to which the statement is committed only to facts about the representative content or to facts about the intentional objects.[1]

The distinction between the representative content and the intentional object is parallel to Frege's distinction between sense and reference. Just as a definite description refers to an object in virtue of its sense, but does not thereby refer to its sense, so an intentional state is directed at an object in virtue of its representative content, but is not thereby directed at its representative content. Both sentences describing intentional states and sentences describing acts of referring are subject

to extensional and intensional interpretations and both for the same reasons. Thus 'John referred to the King of France' and 'John thought about the King of France' have both extensional and intensional readings, depending on whether they are construed as about objects of representations or solely about the representations themselves.

The reason that many authors have failed to see that the intentional object is identical with the actual object is that the specification of the intentional object is by way of the aspect under which it is represented by the representative content, and for that reason it is an intensional-with-an-s specification. To put it crudely, it has seemed to them that the intentional object is incomplete in ways that actual objects are not incomplete and that therefore intentional objects can never be actual objects. How can the actual object be identical with the intentional object when the actual object has all sorts of features the intentional object does not have? Thus Davidson (1968:81-82) writes:

> What is less obvious, at least until we attend to it is that the event whose occurrence makes 'I turned on the light' true cannot be called the object, however intensional, of 'I wanted to turn on the light'. If I turned on the light, then I must have done it at a precise moment, in a particular way--every detail is fixed, but it makes no sense to demand that my want be directed at an action performed at any one moment or done in some unique manner. Any one of an indefinitely large number of actions would satisfy that want, and can be considered equally eligible as its object.

But one might as well argue that a description can never be a description of an actual event because events have all sorts of features not included in the description, and any number of other possible events could have satisfied the description. I want to argue on the contrary that since the specification of a want will specify an intentional object under certain aspects, namely those under which it is wanted, it follows only that the specification of the want is intensional-with-an-s. For example, from the fact that I want E, and E is identical with F, it does not follow that I want F.

But it does not follow from the fact that specifications of wants are intensional that actual events and states of affairs cannot be the objects of wants. To conclude otherwise is to confuse properties of the specification of wants with properties of wants, a common confusion which we will explore in a moment.

If we were to generalize the form of Davidson's argument in a way I am sure he never intended, it would come out as an invalid derivation of a false conclusion from two true premises.

1. An intentional state represents its intentional object under certain aspects only.

2. Actual objects (states of affairs, etc.) have all sorts of features under which they are not represented by intentional states.

3. ∴ Actual objects can never be intentional objects.

But 3. does not follow from 1. and 2. An exactly parallel argument would be

1* A description of an object represents the described object only under certain of its aspects.

2* Actual objects have all sorts of features under which they are not represented by descriptions.

3* ∴ Actual objects can never be described objects.

But why has the first form of argument seemed so plausible in a way that the second has not? And why in general do philosophers say such weird things about intentional objects? A large part of the answer I think is that discussions of intentionality suffer from a confusion of levels which is somewhat analogous to a use-mention confusion. Because sentences about intentional states are (often) intensional-with-an-s, it is tempting to conclude that somehow the intentional states themselves are intensional-with-an-s, that somehow they are not really directed at objects but at their own representative content. But this is to confuse properties of the description of intentional states with properties of the state described, to confuse the 'mention' of the mental state with its 'use'. Specifications of intentional mental states are indeed (in general) intensional-with-an-s because they are about representations. But intentional mental states are not <u>about</u> representations, they <u>are</u> representations. In exactly the same way the description of a description of an object may be intensional but it does not follow from that, nor is it generally the case, that descriptions of objects are intensional. Just as ground floor descriptions of objects are extensional, so ground froor intentional states are extensional. That is, insofar as it makes sense to apply the terms 'intensional' and 'extensional' to mental states,[2] there is nothing inherently intensional about intentionality. Just as my statement that Caesar was Emperor of Rome is extensional, so my belief that Caesar was Emperor of Rome is extensional and for the same reasons. Of course your statement that I stated that Caesar was Emperor of Rome is intensional; just as your statement that I believe that Caesar was Emperor of Rome is intensional and again for the same reason: the first pair represent a state of affairs, the second pair represent representations of a state of affairs. This confusion between representations and representations of representations is quite pervasive. Thus, it is often said that propositions are intensional entities. But there is nothing intensional about, say, the proposition that Caesar was Emperor of Rome. It is as extensional as it can be. Of course the expression 'The proposition that Caesar was Emperor of Rome' is intensional on both our criteria. But it is a use-mention confusion to confuse features of the expression with features of what the expression is about. Sentences about propositions are intensional and sentences about mental states are intensional but propositions are not in general intensional and mental states are not in general intensional.

In favor of the view that there is something intrinsically intensional about intentionality, it is sometimes pointed out that a man can have an

intentional mental state even though no object or state of affairs satisfies the content of his intentional mental state. A man can expect rain and worship God even if it does not rain and God does not exist. But analogously a man can predict rain and assert the existence of God even if it doesn't rain and God does not exist. Just as his expectation may be unfulfilled and his worship directed at nothing, so his prediction and his assertion may be false. For the case of so-called propositional intentional attitudes what corresponds to the truth conditions of a statement is the state of affairs which satisfies the representative content of the attitude, what I have been calling the intentional object. For the non-propositional cases, like love and hate, what corresponds to the object one refers to in the use of a referring expressions is the intentional object that one's attitude is directed at. And just as the fact that one may make a statement which is false does not show that one's statement is intensional or one may fail to refer does not show that reference is intensional, so the fact that one's intentional states may not have intentional objects does not in the least show that they are intensional.

The picture of intentional states that is emerging from this discussion is this: Every intentional state consists of a representative content in a certain mode. The same content can occur in different modes, as for example, when I believe it will rain, hope it will rain, want it to rain, etc. It is necessary to distinguish the representative content from the intentional object. An intentional state will have an intentional object if and only if its representative content is satisfied by an object or state of affairs. I have so far said nothing about how these representative contents are realized, whether by words, images, dispositions to behavior or neurophysiological states, or what not, and I do not think it matters much for the purposes of the discussion so far. On this account, there is nothing ontologically peculiar about intentional objects; they are just ordinary objects and states of affairs at which our mental states are directed. To say that Sally is the intentional object of Bill's love of Sally is just like saying Sally is the described object of Bill's description of Sally, it assigns no ontological peculiarity to Sally. Intentionality-with-a-t and intentionality-with-an-s are not even remotely in the same line of business. Intentionality-with-a-t is so to speak a ground floor property of the mind. It is how the mind grasps other things. But intensionality-with-an-s is primarily a property of sentences and other forms of representation. Some though not all intensional sentences are about intentionality-with-a-t.

Broadly speaking, we can divide intentional states into two kinds those like belief and desire that represent states of affairs (the so-called propositional attitudes) and those like love and hate that can also represent objects, events, etc. I have so far been speaking of the intentional objects as if they could be either particulars or states of affairs. But that can be misleading. States of affairs are not particulars, nor are they anything like particulars. If my belief that it is raining is correct, then what stands to my belief as 'intentional object' is the same as what stands to my true statement that it is raining as truth conditions. We can if we like call that an object, but to do so encourages a mistake analogous to the mistake of assimilating stating to referring. John's love for Sally has a genuine object, viz., Sally. John's belief that

it is raining will be correct under certain conditions, but those conditions are not an object like Sally, and it does not represent those conditions in the way that 'Sally' represents Sally, but rather in the way that the statement that it is raining represents those conditions. We can continue to talk of intentional objects of propositional attitudes but it is not essential to do so, anymore than it is essential to talk of objects of statements or orders, and if we do so we must avoid confusing these with particulars.

A brief digression about 'desire' and 'want' before we continue: Desires are propositional attitudes; they always have states of affairs as 'intentional objects' and never simply objects. In the formal mode: 'desire' and 'want' are like 'believe' and unlike 'love' and 'hate' in that they must take a sentential complement as a grammatical direct object. There is a very simple syntactical argument in favor of this conclusion. Consider the sentence 'I want your house next summer'. What does 'next summer' modify? It can't be 'want', for the sentence does not mean 'I next summer want your house'. It means 'I now want your house next summer', and the meaning of that can be represented as 'I now want that I have your house next summer'. Since any sentence of the form 'I want x' admits of similar adverbial modifiers, it would seem that wanting requires a propositional representative content.

II

Because there are so many close connections between intentional states and speech acts, it is tempting to suppose that the explanation of intentionality must be linguistic. That is, it is tempting to suppose that the representations which intentional states contain must be themselves sentences or at least part of some languagelike system of representation. It is in this spirit that various philosophers have analyzed belief in terms of dispositions to assert or assent to sentences and that others have proposed that to have an intention to do A is to say to oneself 'I will do A'. On this view (a large part of) the philosophy of mind is a branch of the philosophy of language. And clearly there is much plausibility and indeed some truth to this approach: for all but the simplest intentional states, a being could not even have the state unless he had the linguistic capacity to give expression to the state, as Wittgenstein is constantly reminding us. But it seems to me equally clear that this view cannot be right. Beings without language and without language-like systems of representation can have beliefs, intentions, desires, and expectations. Only those in the grip of a philosophical theory would deny that dogs and small children can, say, desire bones and milk respectively. Furthermore, the direction of explanation seems to be wrong. Among other things what the possession of language enables us to do is give expression to beliefs, desires, and other intentional states. True, given the system of representation provided by language, we can have vastly richer and more complex intentional states than we could without language. But language does not create intentionality; rather, as I shall argue, in an important sense intentionality provides the foundation for linguistic acts. On this view the philosophy of language is a branch of the philosophy of mind.

As a step toward arguing in favor of this view, let us begin by

making explicit several of the analogies and connections between intentional states and linguistic acts.

First, the distinction within the theory of speech acts between propositional content and illocutionary force carries over to intentional states. Just as we can say that my prediction that you will shut the door and my order to you to shut the door both contain the same propositional content, namely that you will shut the door, so we can say that my belief that you will shut the door and my desire that you should shut the door both contain the same representative content, namely that you will shut the door. Furthermore, the state of affairs that makes the statement true or the order obeyed is exactly the same state of affairs that makes the belief correct and the desire fulfilled. Let us introduce the notion of satisfaction to describe what is common to the notions of an order being obeyed, a statement being true, a desire being fulfilled, and expectation coming to pass, etc.: a mental state or linguistic act is satisfied if and only if the state of affairs specified in the propositional content actually exists. We can then say that the conditions of satisfaction of speech acts are closely paralleled by the conditions of satisfaction of intentional states. More about this parallelism in a moment.

Secondly, the distinction in direction of fit that applies to speech acts also applies to intentional states. The distinction in direction of fit between an order and a statement can be characterized by saying that in the case of an order, it is the responsibility of the person ordered to make his behavior match the propositional content of the order (and hence orders have the world-to-word direction of fit);[3] whereas in the case of a statement, it is the responsibility of the speaker to make the propositional content of his statement match the world (and hence statements have the word-to-world direction of fit). Apologies, thanks, and congratulations have no direction of fit because in their performance the speaker takes for granted that the propositional content is already satisfied. When I apologize for stepping on your toe, I neither assert that your toe has been stepped on, nor do I try to get your toe stepped on. Rather, I presuppose that it has been stepped on. These distinctions carry over exactly to intentional states. Belief, one might say, has the mind-to-world direction of fit, because if my belief is mistaken it is the belief and not the world which is at fault. Desire has the world-to-mind direction of fit because if my desire is unfulfilled it is so to speak the world and not the desire which is at fault. Sorrow, gratitude, and pleasure have no direction of fit, rather each of them contains a corresponding belief that the propositional content is satisfied. Thus in order to be sorry for stepping on your toe I must believe that I did in fact step on your toe.

A third connection between intentional states and linguistic acts underlies the first two. In general, a speech act with a propositional content is an expression of the corresponding intentional state, and the propositional content of the speech act is identical with the representative content of the intentional state. Thus when one states that p one expresses (whether sincerely or insincerely) the belief that p, when one orders that H does A one expresses a wish or desire that H does A, when one promises to do A one expresses an intention to do A, when one apologizes, thanks, or congratulates one expresses sorrow, gratitude, or

pleasure about the states of affairs for which one is apologizing, thanking, or congratulating. Furthermore, the expression of the intentional state is not a mere accompaniment: there is an internal connection in the strict sense between the performance of the speech act and the expression of the corresponding psychological state, as is shown by Moore's paradox. One cannot say, 'It's raining but I don't believe it's raining', 'I order you to leave but I don't want you to leave', 'I promise to come but I don't intend to come', 'I apologize but I am not sorry', 'thanks for giving me the money, but I am not glad you gave me the money'. In each case the absurdity derives from the fact that one is denying the presence of the intentional state which was expressed in the immediately preceding portion of the utterance. Of course one may express an intentional state one does not have, and that is how insincerity in speech acts is possible. Indeed, each of these intentional states is the sincerity condition of the corresponding speech act, because its presence or absence determines what constitutes sincerity or insincerity in the performance of the speech act. When one makes a statement, for example, one necessarily expresses a belief but one does not necessarily have the belief one expresses.

These three connections show the close parallelism of speech acts and intentional states. Both speech acts and intentional states present a propositional content in a certain mode, a psychological mode in the case of intentional states and an illocutionary mode in the case of speech acts. But in the performance of the speech act, there are not two things going on, the performance of the act and the expression of the corresponding psychological state. Rather the performance of the act is eo ipso, an expression of the corresponding intentional state and the propositional content of act and state are identical. Where there is a direction of fit, the speech act will be satisfied if and only if the intentional state is satisfied. Thus the order that p will be obeyed if and only if the desire that p is fulfilled, the statement that p will be true if and only if the belief that p is correct, the promise that p will be kept if and only if the intention that p is carried out, etc.

Intentionality-with-a-t thus infects nearly all uses of language and this fact requires explanation. A tempting but obviously false explanation is to say that speech acts just are expressions of psychological states and that is all there is to it. While this explanation has some plausibility for the expressive class of speech acts--such as apologizing, thanking, congratulating, etc. it is obviously false for statements, orders, promises, and a large number of others, false because there is more to a statement than just an expression of belief, more to a promise than just an expression of intention, and more to an order than just an expression of a desire. But what more is there? And what is the relation between this 'more' and the underlying intentional states?

As a way of getting at the answers to these questions, imagine a class of beings who were capable of having intentional states like belief, desire, and intention but who did not have a language. What more would they require in order to be able to perform linguistic acts? Notice that there is nothing fanciful in the supposition of beings in such a state, since as far as we know the human species once was in that state. Notice

also that the question is conceptual and not historical or genetic. I am not asking what additions would need to be made to their brains or how language did evolve in the history of the human race.

When we have ascribed to our beings the capacity for having intentional states we have already ascribed to them the capacity for relating their intentional states to objects and states of affairs in the world. The reason for this is that a being capable of having intentional states must be capable of an awareness of the conditions under which its intentional states would be satisfied. For example, a being capable of having desires must be capable of an awareness of the satisfaction or frustration of its desires, and a being capable of intentions must be capable of recognizing the fulfillment or frustration of its intentions. And this can be generalized: For any intentional state, a being that has that state must be able to distinguish the satisfaction from the frustration of that state.[4] This follows from the fact that an intentional state is a representation of the conditions of its satisfaction. This does not mean that such beings will always or even most of the time get it right, that they won't make mistakes; rather, it means that they must have an understanding of what it would be to get it right.

Now back to our question: what more would such beings have to have in order to have a language? The question needs to be made narrower, because there are all sorts of features of actual languages that are irrelevant for our present discussion. Presumably such beings would need a recursive device capable of generating an infinite number of representations, they would need quantifiers, logical connectives, modal and deontic operators, tenses, color words, etc. The question I am asking is much narrower. What would they need in order to get from having intentional states to performing illocutionary acts?

The first thing that our beings would need to perform illocutionary acts is some means for externalizing, for making publicly recognizable to others, the expressions of their intentional states. A being that can do that on purpose, that is a being that does not just express its intentional states but performs acts for the purpose of letting others know its intentional states, already has a primitive form of a speech act. But it still has nothing as rich as our notions of a statement, a request, or a promise, etc. A man who makes a statement does more than let it be known that he believes something, a man who makes a request does more than let it be known that he wants something, a man who makes a promise does more than let it be known that he intends something. But again, what more? The primary extra linguistic purpose of having the institution of assertion is to give information, the main purpose of having the institution of requesting (ordering commanding, etc.) is to get people to do things, and the main purpose of having the institution of promising is to create stable expectations of people's behavior. These facts will, I think, provide a clue to the relations between the speech acts and the corresponding intentional states. As a preliminary formulation one might say that our beings would be capable of making a primitive form of assertion when they could perform actions which were expressions of belief for the purpose of giving information, 'requests' (in this primitive form) would be expressions of desire for the purpose of getting other people to do things, 'promises'

(again in primitive form) would be expressions of intention for the purpose of creating expectations in others.

The next step would be to introduce conventional procedures for doing each of these things. However, there is no way that these extra linguistic purposes can be realized by a conventional procedure. They all have to do with the effects which our actions have on our audiences, and there is no way that a conventional procedure can guarantee that such effects will be achieved. The perlocutionary effects of our utterances cannot be included in the conventions for the use of the device uttered, because an effect which is achieved by convention cannot include the subsequent responses and behavior of our audiences. What the conventional procedures can capture is so to speak the illocutionary analogue of these various perlocutionary aims. Thus for example, any conventional device for indicating that the utterance is to have the force of a statement (for example, the indicative mood) will be one which by convention commits the speaker to the existence of the state of affairs specified in the propositional content. Its utterance therefore provides the hearer with a reason for believing that proposition and expresses a belief by the speaker of that proposition. Any conventional device for indicating that the utterance is to have the force of a directive (request, order, command, etc.) will be one which by convention counts as an attempt by the speaker to get the hearer to do the act specified in the propositional content. Its utterance therefore provides a reason for the hearer to do the act and expresses a desire of the speaker that the hearer do the act. Any conventional device for indicating that the utterance is to have the force of a commissive (promise, vow, pledge) counts as an undertaking by the speaker to do the act specified in the propositional content. Its utterance therefore creates a reason for the speaker to do the act, creates a reason for the hearer to expect him to do the act, and expresses an intention by the speaker to do the act.

The steps then necessary to get from the possession of intentional states to the performance of conventionally realized illocutionary acts are, first, the deliberate expression of intentional states for the purpose of letting others know that one has them, second the performance of these acts for the achievement of the extralinguistic aims which illocutionary acts standardly serve, and third the introduction of conventional procedures which conventionalize the illocutionary points that correspond to the various perlocutionary aims.

III

Throughout this paper I have used the concept of representation as an unanalyzed notion. But as we have already noticed statements about representations are intensional-with-an-s and the reason for that is that representation is intentional-with-a-t. So, by describing intentional states using the concept of representation I have described intentionality in intentional terms. Is there any way out of this circle? I do not believe there is. I do not believe there is a nonintentional explanation of intentionality. That is, there is no analysis of intentionality into logically necessary and sufficient conditions of the form 'X is in

intentional state S if and only if p, q, and r,' where 'p, q, and r' make no use of intentional notions.

This does not mean that there is not a great deal more that can be said by way of describing and explaining how intentionality works. The point is rather that such descriptions and explanations will necessarily employ intentional notions. Our characterizations, descriptions and explanations of intentionality will themselves not be reductive, they will presuppose some understanding of intentionality. We might call this pattern of explanation, where intentionality is used in its own characterization, the circle of intentionality or simply, the intentional circle.

It is interesting to observe the forms which efforts to eliminate intentionality take. The standard behavioristic analyses of mental states blandly use the concept of intentional behavior as if it were somehow less mentalistic than the other mental notions, but to say that a man is walking to the store or eating a meal is to attribute to him mental states which are no less mental than to say that he wants to get to the store or he believes that the stuff on his plate is food. We have the illusion that behavior is not a mental notion because we can observe bodily movements, but the bodily movements only constitute human actions given the assumption of the appropriate intentions and beliefs. Thus either the behavioristic analysis assumes intentionality in the analysans, in which case it is another instance of the circle of intentionality; or it doesn't in which case the analysis is inadequate because no behavior, in the sense of human action, is mentioned in the analysis at all. Behavioristic analyses of meaning suffer from this dilemma in an acute form. There have, for example, been efforts to explain semantic notions in terms of the notion of assenting to a sentence. But assenting and dissenting are full-blown illocutionary acts, every bit as rich as promising or stating. To use these notions in the analysis of meaning, or of any other semantic notion, is to presuppose an understanding of speech acts, and since the speech acts concepts contain the semantic notions that the theory was supposed to analyze, the theory is hopelessly circular. Here, the situation is worse than in the circle of intentionality, because the notions of assent and dissent require more than just intentionality, they contain the notion of meaning.

The problem is not, as is sometimes alleged, that sentences asserting that someone assented or dissented are intentional-with-an-s, nor even that assenting and dissenting are themselves intentional-with-a-t, but that assenting and dissenting are both illocutionary acts and thus understanding them presuppose an understanding of the crucial semantic notions, especially meaning.

But to say there is no nonintentional explanation of intentionality is not to say there is no nonlinguistic explanation of language. If we allow ourselves intentional notions in the analysans, then it seems to me, we can give an analysis of the basic linguistic acts in terms which do not themselves employ any semantic notions, but do indeed employ intentional notions. The key to this analysis is that the intentional notions already contain the notion of their own satisfaction, and we can graft our semantic notions onto the intentional notions using this nonsemantic notion of

intentional satisfaction. I have not given that analysis in this paper but have only sketched the direction it might take. It seems to me not at all paradoxical that there should be nonlinguistic analyses of the linguistic but not nonintentional analyses of intentionality. Speaking a language is, after all, a part of human behavior and of human conscious life. It would be surprising if we could not describe it in terms derived from human behavior and human conscious life. But in the way that intentionality underlies the possibility of linguistic acts there is nothing that conceptually underlies intentionality. Intentionality is precisely that feature of mental states, human or otherwise, that enables those states to represent other things.

NOTES

[1]It is a mistake to say that there are different senses of 'belief', a 'notional' and a 'relational' sense. Rather there are different kinds of reports of belief. One and the same belief can be reported in several different ways, e.g., 'John believes truly that fa', 'John believes of a that fa', and 'John believes that fa'.

[2]It is fairly easy to extend our criteria of intensionality to apply to intentional states. If the satisfaction of the state requires the existence of the objects represented and if the satisfaction value of the state is not altered by the substitution of other representations of the same object, then the state is extensional.

[3]For an explanation of the notion of direction of fit, see Searle (1975).

[4]This looks like it leads to an infinite regress, but it doesn't. The understanding of the conditions of satisfaction of our intentional states is not an additional intentional state.

REFERENCES

Davidson, D. 1968. Actions, reasons, and causes. The Philosophy of action, ed. by A. White, 79-94. Oxford: Oxford University Press.

Mackie, J. L. 1975. Problems of intentionality. Phenomenonlogy and philosophical understanding, ed. by E. Pivcevic. Cambridge: Cambridge University Press.

Searle, J. R. 1975. A taxonomy of illocutionary acts. Minnesota Studies in the Philosophy of Science, Vol. VII, 'Language, mind, and knowledge', 344-69.

Wittgenstein, L. 1967a. Remarks on Frazer's Golden Bough. Synthese 17. 233-53.

------. 1967b. Zettel. Ed. by B. E. M. Anscombe and G. H. von Wright; trans. by G. E. M. Anscombe. Berkeley: University of California Press.

ON SURFACE FORM AND LOGICAL FORM

Edward L. Keenan
University of California at Los Angeles &
The University of Tel Aviv

The body of this paper, sections 2 and 3, presents two quite general correlations between the surface forms (SFs) of natural languages and their logical forms (LFs). Section 1 discusses briefly the nature of SFs and LFs, and section 4 assesses the significance for linguistic theory of establishing relations between SFs and LFs.

1. *The Nature of SFs and LFs.* I will assume a SF (= a surface structure) can be represented by a labelled tree (phrase marker) of the sort usual in generative grammar. Subtrees of a tree represent constituents of the SF and are themselves SFs (since they are also labelled trees). The labels on the nodes of a tree represent the grammatical categories and subcategories of expressions of the language.

It is worth emphasizing that the SFs of a language are defined by a complex set of rules which may be empirically validated (judged descriptively adequate) according as (1) the phonologically interpreted SFs they define are judged grammatical (the soundness criterion) and (2) according as all the forms speakers judge grammatical are in fact defined (generated) by the rules (the completeness criterion). In addition a set of rules may be judged explanatorily adequate according as it is useful in explaining other facts about the language, such as that children learn languages quickly on the basis of limited exposure.

Consequently there can be no simple direct argument concerning the correctness of a particular assignment of SF to a particular expression of a language. We can only assess the descriptive and explanatory adequacy of the entire set of rules, and it is this set which determines assignments of SFs in particular cases. Further, as we modify the rules to make them descriptively more adequate our assignments of SF will change (as the history of generative grammar amply demonstrates).

The nonobvious nature of structure assignment is something linguistic theories share with all scientific theories. There is for example nothing obvious, or given, or natural, about the atomic structure which physical theories assign to oxygen. Our representations here are part of a theory which assigns structures to many compounds and elements and whose correctness is determined by its predictions of the chemical and subatomic properties of these elements and compounds. Further, our assignment of structure in particular cases has undergone massive changes over time as our understanding of the chemical and subatomic properties has grown.

What holds of SFs and atomic forms also holds of LFs. Thus given a set of expressions of a language (e.g. English, elementary arithmetic, etc.)

a logical theory defines a set of LFs (logical structures) whose descriptive adequacy is determined by the correct predictions they make concerning the logical properties of the expressions of the language. For example it is a logical property of (1) below that it entails (logically implies) (3), whereas (2) does not have this property.

(1) Both John and Mary can swim

(2) Either John or Mary can swim

(3) Mary can swim

(To say that a sentence S entails a sentence T is just to say, informally, that in any state of affairs in which A is true, B is true. That is, whenever the world is the way A says it is, it is also the way B says it is.)

Furthermore, logical theories for a natural language can also be assessed as more or less explanatorily adequate according as the set of LFs the theory defines is useful in the explanation of other properties of the language. For example, in *Logical Types for Natural Language* (Keenan and Faltz 1978, henceforth *Logical Types*) it is argued on several different grounds that the LFs that theory provides is explanatorily more adequate (in addition to being descriptively more adequate) than those provided in 'The proper treatment of quantification in ordinary English' (Montague 1973). We argue for example that our LFs make fewer demands on the logical competence of speakers (and hence learners) of the language.

As with SFs then there is nothing obvious or given about the assignment of LF to a particular expression of a language. The assignment depends on the descriptive and explanatory adequacy of the logical theory which defines the set of LFs. And as with SFs, as we modify our logical theories for a natural language in the direction of increasing descriptive adequacy the assignment of LFs to expressions of the language will change. For example, a purely sentential logic for English is not rich enough to show that the a-sentences below entail the corresponding b-sentences:

(4) a. All men are mortal and Socrates is a man
    b. Socrates is mortal

(5) a. All horses are animals
    b. All heads of horses are heads of animals

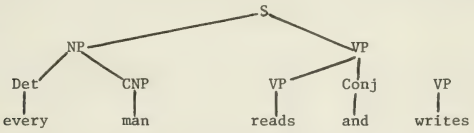
First order logics, which distinguish a predicate argument structure in simple formulas and which allow quantification over arguments can represent the entailment relations given above. And recent work in natural logic, such as Montague Grammar generally (cf. Partee 1976) and *Logical Types* distinguishes more structure than the classically-given first order predicate calculus (CL) and is descriptively more adequate for English than is CL. (Richness of logical structure does not necessarily correlate with being a higher than first order logic. See section 4 for discussion.)

In addition to the theoretical and methodological properties which

scientific theories in general share, syntactic theories and logical theories have some additional properties in common. Namely the structures which both define are representable by labelled trees. That is, they are linguistic objects, having a constituent structure, etc. It thus makes sense to ask to what extent the SF a syntactic theory assigns to an English expression corresponds or 'looks like' the LFs assigned to that expression by a logical theory. Is there for example a one-one correspondence between the constituents of the SF and those of its LFs? Do constituents of different grammatical categories correspond to logical constituents of different logical categories? And so on.

Answers to these questions depend of course on the exact nature of the SFs and the LFs the two theories define. Consider for example the English SF in (6).

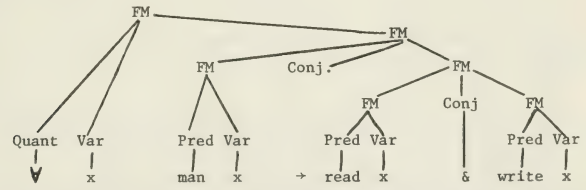
(6)



The example is for illustrative purposes and does not for example represent the nonsegmented Auxiliary or the verb agreements. Further we do not insist on the label CNP, common noun phrase, but we do insist that *every student* and *student* in (6) be assigned different category labels, a distinction recognized both by traditional grammar and X-bar notation. We shall henceforth for clarity refer to expressions like *student*, *fat student*, etc., as CNPs and expressions like *every student*, *a student*, *John*, *he*, etc. as DNPs, determined noun phrases.

A plausible LF for (6) in CL would be:

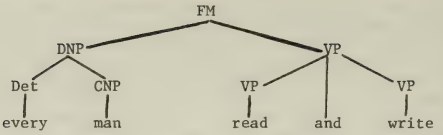
(7)



Clearly the correspondence between (7) and (6) is none too good. (7) for example contains five properly embedded formulas (FMs), the category which corresponds to Sentence, whereas (6) has no properly embedded sentences. Further the category distinction in English between CNP and VP is lost in CL, both being represented as one-place Predicate symbols.

By contrast the LF for (6) given in *Logical Types* is:

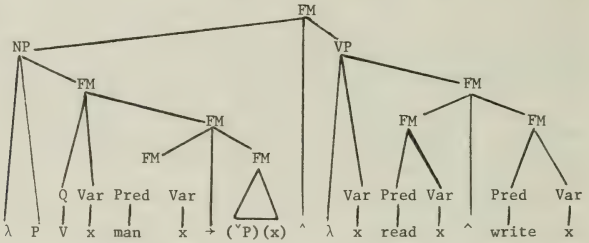
(8)



Obviously the correspondence between constituent structures is close to one-one here, and all the category distinctions in English are preserved in the LF. *Logical types* does not represent Auxiliaries or verb agreements however and further the trees are unoriented. That is, left-right order is not represented in (8). (8) is in fact a set of two expressions, the DNP *every man* and the VP *read and write*. The DNP is itself a set of two expressions whose members are *every* and *man*, and the VP is a set of three expressions: *read*, *and*, and *write*.

On the other hand, (9) below would be, if I am not mistaken, a LF assigned to (6) in Montague Grammar. (I have used where possible 'standard' names for categories rather than those actually used in MG.)

(9)



Clearly the correspondence here is less good than in CL. Not only do we lose the distinction between CNP and verb but (9) presents seven properly embedded formulas.

Given a set of SFs and a set of LFs for a natural language we may now say that a property P defined on the SFs corresponds to a property P' defined on the LFs just in case a SF has P iff some (all) of its LFs have P'. And in general SFs will correspond to LFs to the extent that their properties correspond. Further, this notion of correspondence can be naturally generalized in several ways. We may replace P and P' by n-place relations R and R' and we may allow the arguments of R and R' to be not simply SFs and LFs themselves but sets of SFs and LFs.

To support our claim that in general there is a significant correspondence between the SFs for a natural language and their LFs we shall consider here two relations defined on SFs, the agreement relation and the left-right order relation and argue that in a large class of syntactically specifiable contexts these relations do correspond to relations defined on LFs.

2. *The Agreement Relation.* We observe across a wide range of languages that certain constituents of SFs 'agree' with certain others. For example in a SF for a sentence consisting of a VP and a DNP (regardless of their relative order) the VP may agree with the DNP. That is, a given expression of the VP will vary in form according as expressions for the DNP are drawn from different gender classes and have different numbers. In the same sense we find that transitive verbs (TVs) and in fact transitive verb phrases (TVPs) sometimes agree with their DNP objects (the DNP they combine with in whatever way to form a VP), though this agreement is less common than VP agreement with subjects. Another common case is that of adjective, and more generally adjective phrase (AP) agreement with their CNP 'heads'.

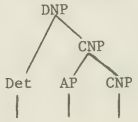
But we don't expect that just any constituent of a SF may agree with any other. It is then an interesting problem in Universal Grammar to specify the pairs of SFs <A,B> such that A may agree with B in a syntactic structure E. (By may agree I mean that there are possible human languages in which they do agree.)

To define the problem somewhat more precisely, let us consider here a syntactic structure (and so in particular a SF) to be a labelled tree of the usual sort less the terminal symbols, that is, less the actual expressions of the language. We may then say of two expressions in a language that they have, or express, a same SF. Arguably for example both *some tall man* and *every woman* express (10a) below, though only *some tall man* expresses (10b).

(10) a.



b.



We may now define:

(11) Given A and B distinct constituents of a syntactic structure E in a language L, A agrees with B iff the form of expressions of A varies with the choice of expressions of B.

(To say that A and B are distinct is to say that neither is a subconstituent of the other.) And our problem may now be stated thus:

(12) For which choices of E, A, and B as above is it the case that there exist languages in which A agrees with B?

A reasonable starting point is the common cases of agreement cited above: VPs may agree with their DNP subjects; TVPs with their DNP objects; and APs with their CNP heads. Do these cases have anything in common? I think so. Namely, there is a sense in which how we semantically interpret the item which agrees varies with how we interpret the item it agrees with. This claim is not obvious and will be justified below. Specifically, let us consider the Meaning-Form Dependency Principle as a first attempt to characterize the class of possible agreement phenomena.

THE MEANING-FORM DEPENDENCY PRINCIPLE (MFDP)

> Given A and B distinct constituents of a syntactic structure E, A may agree with B iff the semantic interpretation of expressions of A varies with the semantic interpretation of expressions of B in the interpretation of E.

Loosely, the form of A's varies with that of B's just in case the interpretation of A's varies with that of B's. Which is to say that a surface relation holds of A,B pairs just in case a certain semantic relation holds between their interpretations. And this is just the kind of correspondence we are looking for, although the MFDP does not specifically mention a relation defined on LFs to which the agreement relation on SFs corresponds. We will state such a relation shortly, but as it is the MFDP which gives the intuition behind the principle, let us first see what justification there is for it.

Consider first the case Adjective Phrases. If the principle is correct it should be the case that how we interpret modifying adjectives varies with the interpretation of the CNP they modify. And this is in fact the case. Consider for example an ordinary adjective like *flat*. In expressions like *a flat road*, *a flat table*, etc. *flat* means something like 'without bumps or depressions'. And this in general seems to be the interpretation of *flat* when the CNP it modifies refers to a class of objects with extended surfaces. But if the CNP refers to a normally tasty liquid, as in *flat beer/champagne* then *flat* means something like 'lacking its normal taste, having lost its "zip"'. And if the CNP refers to a class of objects which, in their usable state, are inflated with a gas of some sort, as in *a flat tire/balloon* then *flat* means something like 'deflated'. And if the CNP refers singing or voice quality, as in *a flat voice*, *flat sopranos*, then *flat* means something like 'too low in pitch'. In short, our interpretation of *flat* is adjusted to, or varies with, the kind of thing we are predicating it of.

And this is commonly the case for common adjectives. Thus strong tables or fortresses are ones which can withstand much force, whereas strong wrestlers or weightlifters are ones that can exert much force; and strong tea or tobacco are substances which produce a marked effect on those who ingest them. Similarly *solid* buildings or tables are basically strong ones; but solid bars or cubes are ones which are not hollow, solid gold or silver are substances of a high degree of chemical purity, and solid lines or rows of buildings are ones without a break.

We may infer then that how we interpret modifying adjectives, and thus APs in general, varies according to the meaning of the CNP they modify.

Analogous claims hold for basic transitive verbs. Consider for example *cut*. If its object is (roughly) an animal or a largish body part of an animal, *cut* means something like 'to make an incision in the surface of'. Note that the integrity of the object cut is understood to be preserved. A finger which has been cut is still a finger, and can be cut again. Thus *cut* when applied to animals or body parts does not imply cutting all the way through. Nor does it imply intentionality or a purpose on the part of the

agent. I may cut John either accidentally or on purpose, and in each case the adverbial phrases *accidentally*, *on purpose* add information not present in the meaning of *cut* itself. On the other hand, if the object of *cut* is a prepared foodstuff, as in *to cut a cake*, *the roast*, *cut* is now interpreted as something like 'to divide into portions for the purpose of serving'. So the action *cut* specifies in these cases is an intentional or purposive one. Further, the integrity of the object cut is seriously affected and normally destroyed as we do now cut all the way through. On the other hand, if the object of *cut* refers to the kind of object consisting of many uniformly elongated objects, as in *to cut the lawn*, *your hair*, *one's fingernails*, *cut* is interpreted to mean something like 'trim'. In distinction to the first sense of *cut* above (cut a person), we do cut all the way through in that parts of the whole are cut off. And in distinction to the second sense above (cut the cake) we do preserve the integrity of the whole. Once we have cut the lawn there is still a lawn there. And further, some notion of purpose is implied. We cut the lawn or our hair to make them more regular, pretty, or something like that.

Further, if the object of *cut* is a concrete mass noun, as in *to cut alcohol* or *cut heroin*, *cut* means something like 'diminish the potency of by admixing a physically comparable substance'. Thus we may cut whiskey with water, marijuana with tea, etc. No notion of making an incision is present in any sense in this interpretation of *cut*. And if the object of *cut* is an abstract one of certain sorts, as in *to cut working hours/production quotas/prices*, *cut* means something like 'to decrease the value of along a numerically continuous dimension (the dimension in question of course is one appropriate to the nature of the object cut). And matters are still more varied. Thus *to cut a class or a meeting* means to not attend when supposed to; *to cut a path* (through a field) or *a tunnel* (through a mountain) means to create a path or a tunnel by cutting something else and removing it; *to cut a film* or *a dissertation proposal* means to eliminate parts of it (destroying integrity) and put the result back together again (restoring integrity) yielding a shorter and presumably better work (notion of purpose involved).

In short then, the action we interpret *cut* as specifying is one appropriate to the object affected. And what holds of *cut* here holds as well of other common transitive verbs. To *drive cars or buses* is a different action from driving cattle or sheep; and both actions are different from driving a motor or a generator. *To drill* means something different in *to drill recruits* than in *to drill metal*, etc. Thus transitive verbs, and so transitive verb phrases in general, vary in their interpretation with the interpretation of the object affected.

Finally, intransitive verbs, and so VPs in general, behave similarly. Thus to say of an animal that it *is running* is to say both that the animal is moving its parts (internal movement) and that it is changing location however momentarily relative to other things (external movement). But to say that a watch or a car motor is running is at best only to make a claim concerning internal movement; no external movement is implied. If you are running with your watch it does not follow that your watch is running. And to say that your nose or the faucet is running implies neither internal nor external movement on the part of the nose or the faucet. Rather it

implies that liquid, of a sort appropriate to the referent of the subject phrase, is moving out of that referent. And to say of a play or an exhibition that it is still running is to say that it is still available for public enjoyment, with no implication of physical movement at all. In short, how we interpret *run* in a context 'DNP is still running' depends on what sort of object the DNP refers to.

Note that the dependency here is clearly semantic. If we say *it is still running* how we interpret the verb depends on what *it* refers to. And if I ask you to *cut them* the action I'm asking you to perform is different according as *them* refers to production quotas or salamis. And the interpretation of flat in *a flat one* or *flat ones* depends on the class of item referred to by *one* or *ones*.

It does appear then that the MFDP holds for the common cases of agreement considered. But many questions concerning this dependency still need to be answered. At what level of linguistic description should this meaning dependency be represented? Semantic? Pragmatic? (We have remained noncommittal on this point.) Does the same sort of form-meaning dependency obtain for other cases of agreement phenomena? And is there any correlation between what varies in meaning with what and the logical structures assigned to agreement pairs?

Let us consider the last question first. Within Classical Logic there seems no complete description of the agreement pairs (VP and NP, TVP and NP, AP and CNP). Simple VPs correspond to one-place predicate symbols, and a limited class of DNPs, such as proper nouns, corresponds approximately to individual constants (0-place function symbols), so we may consider that simple cases of [$_S$DNP, VP] structures correspond to atomic formulas, in which, semantically, the VP is interpreted as a function taking the DNP interpretation as an argument. But the correspondence breaks down when more complex DNPs in English are considered. Thus *every man is running* does not correspond to a formula with a function-argument structure even grossly isomorphic to that of *John is running*. The two cases may be compared as follows:

(13) a. John is running
b. R(j)

(14) a. Every man is running
b. $(\forall x)(M(x) \rightarrow R(x)$

In fact, nothing in (14) corresponds to (interprets) the English DNP *every man* at all.

Transitive verbs correspond somewhat to two-place predicate symbols in CL, but a VP in English consisting of a transitive verb plus object does not correspond to any logical category in CL. On the other hand, it is easy to reformulate CL without changing the entailments of any of the formulas in such a way that transitive verbs do correspond to something which combines with a simple DNP to form a VP. A classical two-place predicate is semantically interpreted as a function U^2, the set of ordered pairs on U, the universe of discourse (i.e. the entities which exist in

that state of affairs) into the set of truth values, represented as 0 (false) and 1 (true). Given such a function f, there is a unique other function g which maps members of U onto VP meanings, that is functions from U into {0,1} such that for all x,y in U, $(g(x))(y) = f(x,y)$. So CL could be trivially reformulated so that, for a very simple class of DNPs, transitive verbs in English would correspond to functions taking DNP meanings as arguments, and thus, for this limited class of DNPs, there would be a similarity between VP and TVP representations in CL. Namely, both would be functions taking DNP meanings as arguments.

But when the AP,CNP pairs are considered CL has nothing to offer us. Expressions like *flat road* are not directly assigned any logical category. Nor is there anything in CL equivalent to a modifying adjective. The best translation in CL of (15) below would be (16).

(15) Every happy man is laughing

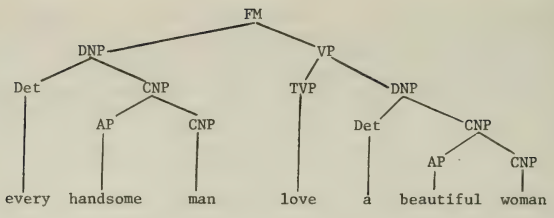
(16) $(\forall x)(man(x) \ \& \ happy(x) \rightarrow laugh(x))$

So the correspondents of 'happy' and 'man' do not form a logical constituent, and are in fact of the same logical category (one-place predicate), and that category is the same as that for simple VPs (*laugh*).

Montague Grammar might appear to offer a more promising set of logical structures since at least the formal English (which is 'translated' into the formulas of an intensional logic and only then interpreted) does have categories of AP and CNP. However, even ignoring the fact that the distinction between CNP and VP does not survive the translation into the set of logical structures, I can still find no basis for a similarity between the three agreement pairs we have considered. It is the case that APs and CNPs form logical constituents, and the APs are interpreted as functions taking CNP meanings as arguments. And it is also the case that TVPs and DNPs form a constituent in the logic and TVPs are interpreted as functions taking DNP meanings as arguments. So for these cases we can say that English item which shows agreement is in each case interpreted as a function taking as argument the (interpretation of) the item it agrees with. But the similarity breaks down when the NP,VP pairs are considered. For here, while such pairs are a logical constituent (formulas), it is not the case that the agreeing item, the VP, is interpreted as a function taking as argument the item it agrees with. Rather the DNP is interpreted as a function taking the VP meanings as arguments.

On the other hand, in *Logical Types* there is a uniform function-argument assignment. VPs are interpreted as functions on their DNP Subjects, TVPs as functions on their DNP objects, and APs as functions on their CNP arguments. Further, the function-argument assignment holds regardless of whether the NPs in question are simple ones like proper nouns, or complex ones like *every flat road*. Thus the LF assigned on that theory to *every handsome man loves a beautiful woman* is:

(17)



We should stress here that (17) is itself a LF in *Logical Types*; that is, it is directly interpreted, not mapped onto a formula in some intensional logic, that is, a different structure, and then interpreted.

Assuming the LFs defined in *Logical Types* we may directly characterize the correspondence between agreement pairs and LFs by:

THE FUNCTIONAL DEPENDENCY PRINCIPLE (FDP)

> Given A and B distinct constituents of a SF E, A may agree with B iff in the LF of expressions of E, the LFs of expressions of A are interpreted as functions taking the interpretations of expressions of B as arguments.

In such cases we shall refer to expressions of A as functional expressions and expressions of B as argument expressions.

Further, as *Logical Types* defines a rather large number of function-argument structures involving DNPs and CNPs, we have at hand a ready set of predictions concerning possible agreement phenomena involving these categories. Let us note the following assignments given in *Logical Types*:

(18) Functions taking DNP arguments:
- a. VPs
- b. TVPs
- c. Possessive Phrases (e.g. *the father of*, *'s father*)
- d. Prepositions

(19) Functions taking CNP arguments
- a. APs
- b. Relative Clauses
- c. Articles (*the*, *a*)
- d. Quantifiers (*every*, *no*)
- e. Numerals (*one*, *two*, etc.)

Considering the second category first, the FDP predicts the existence of languages in which APs agree with CNPs, which we have already seen to be correct. Second, it predicts agreement between Relative Clauses and their CNP heads. This is also correct, for recall the definition of agreement. A agrees with B iff the form of expressions of A varies with the choice of expressions of B. Now a great many languages form Relatives by presenting

in the position relativized a personal pronoun agreeing in noun class with the CNP head of the Relative, as illustrated below from Hebrew:

(20) a. *ha- ish* [*she- natati* *lo* *et ha- sefer*]
RC
the-man that-I+gave to-him DO the-book
'the man that I gave the book to'
b. *ha- isha* [*she- natati* *la* *et ha- sefer*]
RC
the-woman that I+gave to-her DO the book

Clearly the forms of the Relative Clauses vary (one contains *lo* where the other has *la*) according to the gender class of the head CNP. Similarly languages which use special forms of pronouns in Relatives, so called relative pronouns, often present cases where the relative pronoun agrees in gender and number with the head CNP, as thus the form of the entire RC varies with certain choices of expressions for the CNP. So the prediction is verified. Note however, that the FDP does not predict in this case which part of the Relative Clause will vary in form but only that the Relative Clause as a whole will vary.

Third, the FDP predicts that articles may agree with the CNPs they are functions on, and such cases (e.g. Romance) are well known, though perhaps less widespread than the accessibility of confirming cases would indicate.

Fourth, the FDP predicts that quantifer words like *every* and *no* will agree with the CNPs they combine with, and again confirming instances are easy to come by:

(21) a. *tous les hommes* b. *toutes les femmes*
all the men all the women

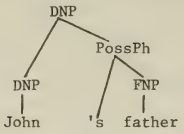
And finally, the FDP predicts that Numerals may agree with their CNPs. Again this is correct; though it is very common that the internal structure of numerical DNPs is complex and nonuniform from one number to another. Nonetheless Russian provides a confirming instance. The numeral 'one' has four forms, one plural form (*odnix*, used before inherently plural CNPs of the sort *pants*, *scissors*, etc.) and three singular forms, one for each of the three genders of CNPs. The amount of agreement decreases rapidly as we consider the higher numbers, but still 'two' in the singular makes at least two gender distinctions.

Recalling now that the motivation for the function-argument assignment given in *Logical Types* (or any logic) is semantic--that is, given that function-argument assignment we get many right and few wrong predictions of entailments, the fact that all the predictions in (19) are correct is rather striking confirmation of the FDP.

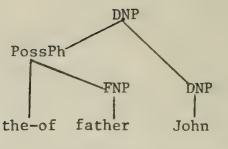
Consider now (18). The FDP predicts, correctly as we have already seen, that VPs may agree with their Subjects, and TVPs with their Objects. Further, within the logic, heads of possessive constructions are assigned a logical category and are semantically interpreted as functions on the Possessor NP. E.g. *John's father* and *the father of John* would be assigned

the following LFs:

(22) a.



b.



(FNP abbreviates Function Noun Phrase, a primitive category in the system.)

And the FDP predicts, correctly, that heads of possessive constructions, Possessive Phrases in *Logical Types*, may agree with their possessors. This in fact is massively common in Amerindian languages, and is evidenced in many others as well, e.g. Turkish, Arosi (Melanesian), Daga (New Guinea), and Hebrew. (23) below from Hebrew is illustrative:

(23) a. *beit-o shel Dan*
house-his of Dan
'Dan's house'

b. *beit-a shel Miriam*
house-her of Mary
'Mary's house'

It is perhaps fair to mention here that while the form of possessive agreement cited above is very widespread--my impression is that it is as widespread as is adjective agreement with CNPs and TVP agreement with their NP Objects--there are some limited cases in which the agreement goes the other way. One obvious case are the pronominal possessives in Romance. Thus we have *son frère* 'his/her brother' and *sa soeur* 'his/her sister'. The possessives here behave in fact like adjectives as regards agreement, and are traditionally called possessive adjectives. The paradigm however is limited to the pronominal possessive forms. The productive possessive construction shows no such agreement, e.g. *le frère de Jean* 'the brother of John' and *le frère de Marie* 'the brother of Mary'. A somewhat larger class of 'wrong' agreements can be found in certain Slavic languages where a DNP with a possessor reading may, in certain cases, be constructed with an adjectival ending and thus agree with the head CNP as an adjective would. Comrie and Thompson (1978) cite the following from Czech:

(24) a. *vědcova kniha*
scientist-f.sg. book
'the scientist's book'

b. *?kniha vědce*
book scientist-gen
'the scientist's book'

Note that (24a) clearly has a possessive meaning, not something like 'a scientific book' (which would be *vědecká kniha*). And contra the FDP the logical possessor shows agreement with the CNP head. Again however the productivity of this paradigm is limited, in fact listable (though the list would be much longer than for the Romance cases cited above). Thus only definite, singular, unmodified possessors can be construed as adjectival. To say *the old scientist's book*, *the (many) scientists' books*, or *a scientist's book* we could only use the 'proper' possessive construction illustrated in (24b). Bob Rothstein (personal communication) points out however that at least certain proper noun possessives, as *Newton's theories*, may also be constructed with the possessor as an adjective.

It would seem then that to maintain the FDP as stated we should weaken it to allow for listable exceptions. And this is not unreasonable. If we establish a correlation between the most productively generated SFs and their LFs we have surely established a correlation between SF and LF in general. Exceptions to the general pattern will have to be learned as special cases. Indeed for other types of correlation between SF and LF, such as some version of the Fregean Principle (meaning of the whole is a function of the meaning of the parts) various types of fixed expressions--proverbs, idioms, etc.--will clearly be listable exceptions.

Finally, FDP predicts that prepositions may agree with the NPs they govern, although the LFs in *Logical Types* are less clear here, as Prepositions have multiple categories.[1] Thus in addition to combining with DNPs to form various classes of modifiers they may also combine directly with VPs and TVPs to extend the class of DNP arguments they take. Thus in *Logical Types* (*sleep*, *in*) is a derived TVP requiring a DNP object to make a VP again (so [(*sleep*, *in*), (*the bed*)] is a derived VP). Further, cases of Preposition agreement with DNPs are harder to find across languages, but some cases do exist. (25) below from Arosi (Melanesian; see Capell 1971 for supporting details) is one example:

(25) *'ini- a mada*
with-it club
'with a club (he hit me)'

Worth noting here is that not all prepositions show agreement with the NP they govern, and those that do, show an agreement paradigm similar (perhaps identical) with that that transitive verbs show with their Objects. However, given the similarity in function-argument structure between transitive verbs and their Objects and Prepositions and their Objects it is perhaps not surprising that transitive verbs can historically become reduced to function as dependent forms, i.e. Prepositions. Note as well that another very common source for prepositions is possessive constructions. Thus for example *behind it* will in many languages be something like *at its back*. If the language were such as to have agreement between possessive heads and Possessors we could expect that prepositions so derived historically might show such agreement as well. Something like this may be the explanation for the postposition agreement in Daga (New Guinea; see Murane 1974 for details). In any event there are several clear cases of postpositions agreeing with their DNP Objects, and the agreement suffixes are the same as the intimate possessive agreements with possessors:

(26) *ge orup ame ena- m ak*
you fellow that with-their go+2sg+imp
'(You) Go with those fellows'

Thus for all the function-argument structures in *Logical Types* which concern either DNPs or CNPs, the Functional Dependency Principle correctly predicts the possibility of agreement. And as these nine cases represent a reasonably large sample of the possible agreement phenomena known to me, I feel that even if the FDP were restricted to hold for just these cases it would still represent a very significant correlation between SF and LF and thus support the general claim that there exist correspondences between

SF and LF. On the other hand, a large number of questions concerning the FDP remain. In section 2.1 below we raise some of these questions.

2.1 *Refining the FDP.*

2.1.1 We have replaced the original intuition expressed in the Meaning-Form Dependency Principle by one, the FDP, stated in formal logical terms. It is reasonable to ask how well the original intuition behind the MDFP carries over to the larger range of cases. The results here are mixed and interesting, and suggest some additional correlations between SF and LF.

Consider first the Possessive Phrases. Do the heads of such constructions vary in their interpretation with the reference of the Possessor in anything like the same way in which say *flat* varies with the interpretation of the CNP it modifies or *cut* varies with the semantic nature of its Object? Here there does seem to be some variation, though it is not clear that it is of exactly the same sort as the very idiosyncratic variation exhibited by *flat*, *cut*, and *run*. Thus consider a PossPh like *the middle of*. If its argument, the Possessor, denotes an activity, such as the performance of a play or a radio program, then *the middle of* picks out that part of the performance which occurs at a moment, or period, of time midway between the beginning and the end of the activity. Analogous claims hold for arguments that denote specified periods of time, such as *day*, *night*, etc. But if the argument refers to a concrete physical object *the middle of* has different senses. Thus the middle of the Earth is a physical point, or region, more or less equidistant from all the points on the surface of the Earth. So *the middle of* may have a temporal or physical interpretation according to the nature of the Possessor--a fact which is well known. But this variation is not wholly similar to that evidence for *flat*, *out*, etc. That is, the same variation shows up for a few other such expressions, e.g. *the end of*, *the length of*, whereas I don't expect (but haven't in fact checked) that any other adjective besides *flat* would have exactly the four senses attributed to it earlier. On the other hand, *the middle of* does show further different senses which do seem more idiosyncratic when applied to different sorts of physical objects. Thus if the Possessor refers to something conceived of in two dimensions, *the middle of* may pick out a *point* at the geometrical center, as in *the middle of the stage* or *the middle of a sheet of paper*. But if the two dimensional object is notably elongated, *the middle of* picks out a *line* more or less equidistant from the edges, as in *the middle of the road*. Further, if the object is conceived of in three dimensions, then differences in the interpretation of *the middle of* show up according as the object is solid or one habitually entered by human beings. Thus *the middle of the Earth (the Sun) consists of molten iron (Helium atoms)* means that the space at the geometrical center is filled with molten iron, etc. But *John is sitting in the middle of the room* does not suggest that John is suspended in midair. Rather the middle of the room defines a region of space with reference to the center of the (two dimensional) floor.

In addition, many overtly possessive constructions lend themselves easily to 'metaphorical' interpretations of the head (or PossPh). Thus one might argue that *the head of* in *the head of the animal* is interpreted literally, and picks out a certain body part of the animal. But *the head*

of the table or *the head of the department* force *the head of* to be understood in a nonliteral (metaphorical) way. This sort of metaphor is what Reinhart (1976) calls nonpoetic metaphor and is characterized, in our terms, by a reinterpretation of the function expression given a 'nonordinary' argument expression (the Possessor). But whether we take one sense as basic and the others as metaphorical or not, it is still the case that it is the function expression which is subject to various interpretations depending on the (literal) meaning of the argument expression.

In addition we might note that nominalizations of various other sorts of functional expressions (adjectives, VPs, etc.) commonly take the form of possessive constructions, and some, but not all, of the sense variation in the original may be preserved under the nominalization. Thus *the strength of* has a similar range of senses to *strong* (person, fortress, tobacco, etc.). On the other hand, *the solidity of* loses some of its senses. Thus while *the solidity of the chair* presumably does refer to its capacity for withstanding force, *the solidity of the bar* or *the cube* is unclear in meaning, referring most likely in my opinion not to its quality of being not hollow but to its capacity of withstanding force. And the solidity of the gold or silver cannot be used to refer to its chemical purity but must refer to its capacity to withstand force, if indeed it refers to anything at all. Nor can *the solidity of a line* or a row of buildings naturally refer to their quality of being 'perceptually unbroken'. Note also that *solid* loses its chemical purity sense when used predicatively. *This gold is solid* does not mean it is chemically pure. These observations suggest:

THE FUNCTIONAL DEPTH PRINCIPLES

(i) The meaning range of functional expressions decreases in proportion to their internal function-argument complexity, and
(ii) for f a functional expression and e an argument expression, the meaning range of f(e) is less than or equal to that of e.

Thus (i) says basically that more derived items exhibit less meaning variation than less derived items of the same category. And (ii) says that a derived item exhibits less (or at least not more) meaning variation than the item it was derived from.

Thus (ii) would predict for example that *run slowly* should have a narrower (or at least not greater) meaning range than *run*, given that *slowly* is interpreted as a function on the interpretation of *run*. And this appears to be correct. While we can understand *John/the motor/the faucet is running slowly* using *run* in the same sense as if *slowly* were absent, we cannot say *My Fair Lady* or *the Braque exhibition is running slowly*, so *run* has lost one of its senses.

(i) would predict that VPs consisting of a TVP plus object would exhibit less meaning variation than simple lexical intransitive verbs. And this also seems correct. It is not easy to find TVP - DNP combinations in English which exhibit the gradient, idiosyncratic variation exhibited by *run*, though some variation still does exist. Thus *to cross the river here* is interpreted as an activity if the subject is animate (e.g. *the soldiers*), but it is interpreted as denoting a state or a locative relation with

various choices of inanimate subject, e.g. *the bridge*, *the telephone lines*, etc. The Depth Principle also suggests that a predicate adjective may have fewer senses than its modifying use since, e.g., *be solid* is functionally more complex than *solid*. And at least for *solid* this appears to be correct.

The Depth Principle very likely has something right about it then, though much more work investigating particular cases would be needed to place it on firm ground. Furthermore there are other restrictions on the meaning range of functional expressions besides those covered in the Depth Principle. Thus some lexical adjectives (intransitive verbs, etc.) simply exhibit much less meaning variation than do others. Thus the *Random House College Dictionary* (1973) lists only one sense for *eleemosynary*, two for *perspicacious* (one archaic), and three for *melancholy*, whereas it gives 23 for *flat*, 23 for *strong*, and 26 for *solid*.

A kind of limit case here may be *logical constants*, items like *every* and *not* whose interpretations do not vary from state of affairs to state of affairs. And as *Articles*, *Quantifiers*, and *Numerals* are logical constants in this sense we would perhaps not expect them to exhibit meaning variation. And that appears correct. We shall not here present and substantiate a principle which would predict that however. Prepositions, few in number in any language compared to for example CNPs and VPs, seem to lie somewhere between ordinary 'content' items and logical constants. And it is unclear to me whether we want to say that *with* in *to fill the tub with Harry/with water/with joy* exhibits meaning variation or is simply multiply ambiguous.

So more research is needed here. But it does seem to me likely that some version of the MFDP, suitably restricted to allow for inherent restrictions on the meaning ranges of specific items, can be made to work.

2.1.2 It is reasonable to query whether functional expressions taking items other than DNPs and CNPs as arguments also show agreement. If not, the FDP should be restricted to only cover nominal agreement. But our linguistic knowledge is limited here, so we shall not at the moment impose such a restriction. Further, we can find cases where adverbs and PPs inflect for categories that the VPs (and TVPs) they are functions on also inflect for. In Malagasy for example locative complements of verbs are marked for tense, and their marking must be the same as that of the verb:

(27) a.

n-	*andeha*	{*t-* / **Ø-* / **ho-*}	*any*	*Antsirabe*	*Rabe*
past-go		{past / pres / fut}	there	Antsirabe	Rabe

'Rabe went to Antsirabe'

And in Avar (NE Caucasian; see Anderson 1977) certain adverbial constructions agree with the subject, as do the verbs:

(28) a. *emen roq'o-ve v-us:ana*
father home- m m-returned
'Father returned home'

b. *ebel roq'o-je j-us:anax*
mother home- f f-returned
'Mother returned home'

But in the Avar case it is 'surely' more natural to consider that it is the entire VP *returned home* which is agreeing in gender with the subject. The apparent agreement between the VP and adverb is an artefact of how VP agreement is internally realized and not a case of the form of the adverb varying with the choice of verb, as would be required by the FDP.

The Malagasy case is perhaps less easily explained away. Given sentences like *John works and has always worked very hard*, where plausibly VPs with different tense marking occur within the scope of the adverb, it is reasonable to think that tense marking is a property of verbs, and so the adverbs or PPs which must have the same tense marking may well be instances of agreement as per the FDP. But it is also not implausible to think that tense marking is a property of the whole VP or even the entire sentence, in which case apparent agreement in tense internal to the VP would be a result of spreading of the marking on the VP as a whole and not a case of the form of the PP varying with the choice of verb.

A clear case of feature spreading rather than agreement is given by case marking on modifying adjectives. While it would be tempting to consider, with traditional grammar, that modifying adjectives agree in case with their CNP heads, it seems clear (see 2.1.5 below) that case marking is a process which takes the entire DNP in its domain and that in certain languages the case marking on the DNP spreads onto adjectives and determiners. So agreement in case, as opposed to agreement in gender class, is not a case of the form of one expression varying with the choice of another as the FDP requires.

The cases of feature spreading above neither support nor refute the FDP, but they do in certain cases support the phrasal as opposed to the merely lexical nature of agreement phenomena. Thus the FDP claims that VPs in general, not just lexical VPs (intransitive verbs) may agree with their DNP subjects. The claim is not falsified if only lexical VPs show agreement since variation in form of part of the VP is variation in form of the whole VP, but the FDP still does suggest that other parts of the VP may vary in form since it does not restrict agreement to lexical VPs.

Thus the examples of verb and adverb agreement with subjects cited for Avar above clearly illustrate full VP agreement as opposed to merely lexical VP agreement. (29) and (30) below from English also illustrate full VP agreement.

(29) a. He *is* behaving *him*self
b. She *is* behaving *her*self
c. They *are* behaving *them*selves

(30) a. He *has* lost *his* way
b. She *has* lost *her* way
c. They *have* lost *their* way

Note that the VP agreement is still evidenced in cases where the lexical verb itself does not agree, as in (31) below:

(31) a. They should/might/must/won't/behave *them*selv*es* (*himself/ *oneself)
b. She may/must/should/have lost *her* way (*his way/*their way)
c. She tried/promised (Bill)/wanted/to lose *her* way

Similarly both copular verbs and predicate adjectives and predicate nominals show agreement with Subjects in many languages.

(32) a. John *is a* doctor
b. John and Mary *are* doctor*s*

(33) a. Marie *est* intelligent*e*
b. Marie et Francoise *sont* intelligent*es*
c. Jean et Marie *sont* intelligent*s*

(33b,c) show as well that what the form of the VP varies as a function of is the entire DNP Subject and not just the 'head'. To know whether *intelligent* is feminine plural, (33b), or masculine plural, (33c), we must in principle check the gender of each conjunct of the Subject. The point is perhaps even more obvious in languages in which full verbs agree in gender, as the Hebrew examples below illustrate:

(34) a. *Miriam v-Ruti medabrot*
Mary and Ruth are speaking (f.pl)
b. *Miriam v-Dan medabrim*
Mary and Dan are speaking (m.pl)

Overall then the fundamentally phrasal (rather than merely lexical) nature of the agreement phenomena that the FDP predicts seems to us basically correct. On the other hand, whether the FDP makes correct predictions concerning function argument structures not involving DNPs and CNPs must await further logical and linguistic research.

2.1.3 But what about further function-argument structures which do involve DNPs and CNPs? A few plausible cases do come to mind. Thus we strongly expect that ditransitive verbs like *give* and *hand* will be treated semantically as functions taking, at some level of analysis, their Indirect Objects (IOs) as arguments. Assuming that, the FDP would predict the possibility of ditransitive verb phrase agreement with IOs, and this is in fact correct. E.g., Spanish, Basque, and Daga (New Guinea) all evidence such agreement.

What about the logical functions represented by *and*, *or*, and *not*? In *Logical Types* these operators do directly form derived DNPs, but they are treated as operators of a different sort from those already considered. The main difference is that they also combine directly with Formulas, VPs, TVPs, and APs to form derived members of those categories. Further, *and* and *or* are two-place functions rather than the one-place ones considered so far. Linguistically, relative to the category they combine with, a certain minimal variation is known. That is, many languages, for example,

use different morphemes for *and* according as the category of items conjoined is different. The examples below from Malagasy are illustrative, and Payne (1978) provides large numbers of other examples.

(35) a. *mihinam-bary Rabe ary* (*sy) *misotro taoka Rasoa*
eat- rice Rabe and drink booze Rasoa
'Rabe is eating rice and Rasoa is drinking booze'
b. *mihinam-bary sy* (*ary) *misotro taoka Rabe*
eat- rice and drink booze Rabe
'Rabe is eating rice and drinking booze'
c. *mihinam-bary Rabe sy* (*ary) *Rasoa*
eat- rice Rabe and Rasoa
'Rabe and Rasoa are eating rice'

On the other hand, I know of no language in which the form of *and* varies with the choice of conjuncts within a category. Thus, as far as I know, if a language has an overt conjunction for DNPs at all then it uses the same form for *John and Bill* as it does for *Mary and Susan*.

2.1.4 We have considered a rather large number of function-argument structures and at least those which involve DNPs and CNPs are commonly enough expressed in SF by items which satisfy the agreement predictions of the FDP. But are there cases of agreement in SF which are not expressed by function-argument structures of the right sort? There would appear to be at least one: That of pronominal agreement with full DNP 'antecedents'. As 'bound pronouns' can occur in constituents in which their antecedents do not occur, the pronoun will not in general be represented as a function taking the interpretation of the full DNP as an argument. This at least is not the case for pronominal representation in *Logical Types*, nor is it the case in any of the commonly used versions of LFs for natural languages or formal languages. Rather pronominal co-reference is represented by the use of Variable Binding Operators (VBOs), illustrated in (36) below from Classical Logic and (37) from *Logical Types*.

(36) a. Every student laughed and cried
b. $(\forall x)$ (student(x) $\rightarrow$ laughed(x) and cried(x))

(37) a. Every man loves his mother
b. ((every,man),(λx(x,(love,(x's mother)))))

The use of such VBOs seems 'unnatural' in the sense that they do not seem to correspond to anything in SF. Indeed we count it as an advantage of *Logical Types* that very many SFs using quantified NPs which bind the reference of more than one position can be represented without the use of VBOs. For example, (36a) above would be represented in *Logical Types* by:

(38) ((every,man)(laughed, and, cried))

However in the LF for (37a) there is no way to bind *his* without using the lambda operator. And it is clearly not the case that what corresponds to *his*, namely the bound occurrence of x, is interpreted as a function taking the interpretation of (*every,man*) as an argument. Similarly the position relativized into in the representations for Relative Clauses in *Logical Types* is represented by a bound variable.

The best tentative conclusion I can reach concerning these phenomena is that the kind of agreement we see between pronouns and full NP antecedents is different in kind from that that we have considered so far, and a logical characterization of it should proceed along lines different from those in the FDP. I would propose then that the FDP be restricted so as to require that the distinct constituents A and B of E are further required to be of distinct categories (not merely subcategories of one and the same larger category). Since pronouns and full NPs are all DNPs, the FDP would then make no prediction about their agreement possibilities. An additional principle would be required then to characterize pronominal agreement. The principle is in fact easy to state, though we shall not be concerned with its details or justification here. But roughly, if A and B are of the same major category, then A is represented in LF by a bound variable and B is what binds it (i.e., in *Logical Types*, B is the argument of the structure created by that VBO).

2.1.5 A final potential restriction on the FDP concerns case marking of DNPs. In many languages a DNP Subject may have one case marker, the absolutive, if the main verb of the VP is intransitive, and another marker, the ergative, if the main verb is transitive. So the form of the DNP would vary with that of the VP, in violation of the FDP. Similarly in many languages Experiencer 'Subjects' will be marked like IOs (that is, dative) whereas Subjects of activity VPs will be nominative (or unmarked).

Appearances then might suggest that the FDP should be restricted so as to cover just the traditional agreement cases and not, as Emmon Bach points out to me, traditional cases of government. To impose such a restriction we might for example stipulate that the variation in form mentioned in the FDP be restricted to variation with respect to inherent nominal properties such as number and gender. Even so restricted the principle would correctly account for a large number of specific correspondences between SF and LF and so still support the general conclusion that there exist correspondences between SF and LF.

I doubt however that such a restriction need be imposed. Even a brief look at case marking properties of languages shows, as I shall briefly argue below, that case marking is a sentence-level phenomenon. That is, case marking, despite the fact that it sometimes (see below) shows up morphologically bound to DNPs, is a property of sentences or more exactly a relation between a DNP and the entire sentence in which it occurs, and not a property of DNPs given as a function of the verb in the clause in which they occur.

To support this claim, I shall summarize a very large number of cases in which DNP case marking is independent of the choice of the verb. Thus for a given choice of verb a given DNP may have any of several case markers (CMs), and the conditioning factors appear to be almost any conceivable property of the sentence containing the DNP. (The latter claim is surely too strong, but a quick look at case marking systems reveals an astounding variety of such conditioning factors.)

First, it is obvious that locative cases are not predictable from the verb generally. Thus in the same way that in English we may say *John fell/*

on the sidewalk/near the station/in the bathroom, etc., so in case marking languages in the narrow sense, such as Finnish, a given verb like *fall* may take DNP complements in any of various locative cases. We may then dismiss locative cases, and indeed all the semantically rich oblique cases (e.g. instrumentals) from consideration and concentrate on case marking affecting Subjects and Direct Objects.

But even here there are a massive number of cases where the CM on a DNP may vary with the verb held constant.

First, consider the very widespread use of 'definite DO marking'. Many languages, such as Hebrew, Turkish, Persian, Spanish, Malagasy, etc., present DOs of the same verb in different cases according as the DO is ± definite and/or ± animate, the exact conditioning factors varying from language to language. (39) from Hebrew is illustrative:

(39) a. *ani ohev yeladim*
I like child-pl
'I like children' (no CM on the DO)
b. *ani ohev et ha- yeladim ha- ele*
I like DO def-child-pl def-those
'I like those children'

A second, very widespread phenomenon among case marking languages concerns differential marking according as the DNP is understood to be partitively or wholly affected. Thus the differences discussed in Anderson (1971) as between *chew the meat/chew on the meat*, *shoot Bill/shoot at Bill*, etc., show up in the case marking in many languages. The case of Finnish is well known. DOs of many affirmative transitive verbs may be partitive in case if they are only partially affected by the action (and so the action may be understood as incomplete). If they are wholly affected, the CM may be either 'accusative' or nominative depending on other factors to be discussed shortly (see Timberlake (1975 and 1977) for detailed discussion). Edith Moravcsik (1978) points out furthermore that this distinction exists in a great many languages and includes Subjects of intransitive verbs as well. She cites, for example, (40) from Russian:

(40) a. *kash-a ostalas' na stole*
kasha-nom remained on table
'Kasha remained on the table'
b. *kash- i eshcë ostalos'*
kasha-gen still remained
'Some kasha was still left'

Similarly many of Fillmore's pairs show up in the case marking of heavily case marking languages. (41) below from Hungarian is also taken from Moravcsik (1978):

(41) a. *János beültette a kertet fákkal*
John planted (def) the garden(acc) trees+with
'John planted the garden with trees'
b. *János fákat ültetett a kertbe*
John trees planted(indef) the garden+into
'John planted trees in the garden'

(Note that as per the FDP, the verb in (41a) agrees with the definiteness of the DO, but does not agree with indefinite DOs as in (41b).)

Third, note that in some ergative languages intransitive subjects may sometimes take either the absolutive case or the ergative case, with a difference in meaning: the ergative indicating Subject intention or responsibility, the absolutive indicating accidental action. E. Pomo (Uto-Aztecan; Moravcsik 1978) and Bats (Caucasian; Comrie 1973) are examples. (42) below is from Bats:

(42) a. *so* *wože*
I-abs fell
'I fell (not implied that it was my fault)'
b. *as* *wože*
I-erg fell
'I fell (it was my fault)'

Further, even a bare animacy distinction may occasionally trigger different CMs on transitive Subjects. (43) from Gugu-Yalanji (Australia) is cited from Gary and Keenan (1977).

(43) a. *dingkar-angka* *kaya kunin*
man- erg+anim dog hit
'The man hit the dog'
b. *kalka-bu* *kaya kunin*
spear-erg+inanim dog hit
(= instr)
'The spear hit the dog'

The above cases are all ones in which semantic properties of the marked DNP are reflected in the CM but are not predictable from the verb itself. Some of these properties, such as definiteness, may also have a syntactic characterization, in which case it is not easy to decide whether the conditioning factor is semantic or syntactic. But other cases are more clearly syntactic. Thus, as is well known, many languages may case mark full DNPs and pronouns differently. Thus English does not mark full DNPs but does mark pronouns. And Dyirbal (Dixon 1972) and many other Australian languages may case mark full DNPs on an ergative-absolutive basis but pronouns on a nominative-accusative basis.

Furthermore, if a language distinguishes noun classes and also case marks, the form of the case marking may vary with the noun class. E.g., Russian and Latin case marking varies according as the nominal marked is masculine or feminine, singular or plural, etc.

The above collection of facts shows that case marking is sensitive to a large number of syntactic and semantic properties of the DNP affected by the marking, where these properties are not predictable from the verb. Further case marking will also be sensitive to whether the DNP in question forms a logical constituent with a VP or a TVP, and will need to have information regarding the semantic and/or syntactic subclass of the verb. Yet even with all this information the CM is not completely predictable. It still might, for example, mark the DNP as partitive or not, as in the

Russian cases above, or it might mark it as ergative or absolutive, as in the Bats cases above. In either case the CM itself brings in new information (partial affectedness, purposive activity, etc.). All this rather suggests the in fact quite intuitive view that case marking is both syntactically and logically a relation between a DNP and a VP or TVP. And if case marking is treated as a two-place function taking DNPs and verbs of various sorts as arguments then the FDP simply has no prediction to make, except the correct one that the form of the CM may vary with the nature of the arguments (stative vs. activity VP, definite vs. indefinite DNP, etc.).

Even this rather appealing picture however is very badly oversimplified. Evidence from very many languages shows that case marking must have much information available concerning the structure of the entire sentence in which the marked DNP occurs.

Thus, for example, several different cases are available where the CM on a given DNP varies with properties of other DNPs in the clause. Recall the Finnish cases mentioned above. The Object of many affirmative transitive verbs if not semantically partitive, and so marked by the partitive case, may still be either nominative or 'accusative'. The conditioning factor here is whether the transitive verb presents an overt subject or not. Thus, using the traditional case labels, we note the following examples from Comrie (1978) (see Timberlake 1975 for a thorough discussion):

(44) a. *Maija-Ø söi kala-n*
Maija-nom ate fish-acc
'Maija ate the fish'
b. *Syö kala-Ø !*
eat-imp fish-nom
'Eat the fish!'

Similarly objects of impersonally used (i.e. subjectless) transitive verbs also take nominative objects. On the other hand, if the subject of the (affirmative) transitive verb is present, the nonpartitive object is 'accusative' ('antiergative' in Comrie's terminology, but in any event different from the nominative). So case marking of one DNP then is contingent on the presence or absence of another.

A different type of dependency between DNPs is illustrated in Dalabon (Australia; Silverstein 1976, cited in Comrie 1978): Here case marking is contingent on the 'chain of being' relations between the transitive subject and the object. Thus if the subject is animate (as in (45) below) and the object inanimate the subject is not case marked. But if both are animate the subject takes an ergative marker:

(45) *Buluŋan-Ø ga'manbuniŋ*
my-father he-made-it
'My father made it'

(46) *Buluŋan- yi wudawud ga'nan*
my-father- erg baby he-looks+at-him
'My father is looking at the baby'

A third related type of example is given by case marking in ergative languages when a 'normally' transitive verb is used 'absolutively' as *He wrote and wrote* as opposed to *He wrote the letter*. Some ergative languages cited in Comrie (1973) use an absolutive ending for the absolutive use, and others, like Hindi use the ergative. In the former case then, case marking of the subject is not predictable from the identity of the verb.

It appears then that the domain of case marking must not only include the DNP and the VP or TVP but also other DNPs in the clause. And yet we have only begun to touch upon the other properties of the sentence to which case marking must be made sensitive. Thus in many languages (Hindi, Georgian) case marking is sensitive to Aspect. Subjects of perfective sentences are marked in one way, subjects of imperfectives another. And since the same verb can occur in many different aspects this property of case marking is not predictable from the verb.

Further, it is quite common that case marking of DNPs is different according as the sentence they are in is negative or not. Thus, as is well known, DOs in Russian and Czech may be genitive if the sentence they occur in is negative. This is also true for Finnish and Lithuanian, as Moravcsik (1978) points out, and further the same holds for certain intransitive subjects. Compare (47a,b) from Moravcsik:

(47) a. *proexal avtomobil'* (Russian)
went-by car (nom)
'A car went by'
b. *ne proexalo avtomobilja*
neg went-by car (gen)
'Not a car went by'

Analogous claims again hold for Finnish and Lithuanian. Worse, in Kawaiisu (S. Numic, Uto-Aztecan; see Munro 1976 for details) Subjects of negative sentences are marked (overtly) like DOs, whereas in nonnegative sentences they are not overtly case marked. In fact, Subjects of embedded clauses generally take object-marking. So it may be the case that case marking is not independent of whether the sentence the DNP occurs in is a main clause or an embedded clause. In Wappo (Li and Thompson 1975, cited in Gary and Keenan 1977) the overt case marking on Subjects is simply lost in finite subordinate clauses:

(48) a. *ce kew-i ew toh- ta?*
that man-Subj fish catch-past
'That man caught a fish'
b. *Ah ce kew-Ø ew toh- ta? hatiskhi?*
I that man fish catch-past know
'I know that that man caught a fish'

It appears then that case marking is fully clausal in nature. And on reflection this is not surprising. In very many languages pre- or post-positions not only mark DNPs, they also mark various types of subordinate clauses. Thus subordinate clauses like *on arriving at the station, John went straight to the ticket counter; The meeting began upon John's arrival*, etc., may well be translated in heavily case marking languages by ordinary

morphologically bound case markers. Recall in this connection the widespread use of the 'absolutive' construction in the IE languages. (49) from Ancient Greek below is illustrative (diacritics omitted):

(49) *élthont-a eis ten polin, ho didaskal-os edeixe ton paid-a*
coming-acc into the city, the teacher- nom showed the child-acc
to sokrat-e
the Socrates-dat
'Coming into the city, the teacher showed the child to Socrates'

The nonfinite verb on the (subjectless) subordinate clause is case marked accusative, which indicates that the understood Subject of *coming* is coreferential with the DNP marked accusative in the main clause, namely the child. Had *coming* been marked nominative, it would have been the nominative DNP, *the teacher*, who was coming; and had it been marked dative it would have been *Socrates*. And had it been in the 'absolutive' form, the genitive case (ablative in Latin, Dative in Gothic, Locative in Sanscrit), the understood Subject of *coming* would have had to be different from any of the major participants in the main clause and thus already understood from the prior discourse.

The general point here is that case marking commonly has a variety of clausal and cross clausal functions. The domain of case marking then, even restricted to DNP marking, is properly the entire sentence in which the DNP occurs.

Note finally, that if case marking is a sentence level property we might expect reflexes of it to show up elsewhere than on the DNPs. And this in fact is correct, at least where case marking is understood broadly enough to cover markings on indirect objects and obliques. Recall the Latin paradigms in which 'prepositions' may show up either on the DNP, or on the verb, or both. Thus from a verb like *ferre* 'to carry' we may form derived verbs such as *inferre* 'to carry into', *transferre* 'to carry across', *ex+ferre* 'to carry from', *ad+ferre* 'to carry to', etc. The DNP 'Object' of such verbs may or may not also carry the preposition. Thus Caesar may carry the war into Italy, he may 'incarry' the war into Italy, or he may "incarry' the war Italy. Similar double markings occur in Hungarian and to a lesser extent Modern Russian. (50) below from Hungarian is illustrative.

(50) *Janos ra-Ø- te- tt- e a kalap-ot az asztal-ra*
John on-it-put-past- 3sg3sg the hat- acc the table- on
'John put the hat on the table'

Further, it is in fact quite common to find that that element which marks the relation a DNP bears to the action expressed by the verb may be marked exclusively on the verb. Such cases arise commonly in Caucasian languages, Penutian languages (Chinook, Totonac), Ancient Greek (see Keenan 1976 for examples), Arawakan languages (e.g. Machiguenga; see Gary and Keenan 1977 for examples) and Bantu languages. (51) below from Kinyarwanda (see Kimenyi 1976 for a thorough discussion) is illustrative:

(51) *Yohani y- a- andik-iish-ije-ho ameza ikaramu*
John he-past-write-inst-asp-loc table pen
'John wrote on the table with the pen'

We may conclude then that syntactically speaking case marking is a function taking the entire sentence as its domain and thus the only prediction made by the FDP is the correct one, namely that the form of case markers may vary with the properties of the sentence it applies to.

3. *The Left-Right Order Relation.* The function-argument structures concerning CNPs and DNPs distinguished in *Logical Types* turn out to support, quite unexpectedly, another correlation between SF and LF: Namely, the characteristic left-right order of the constituents which express these structures. Thus:

THE SERIALIZATION PRINCIPLE (SP)

Different functional expressions taking the same class of argument expressions tend to serialize on the same side of their argument expressions.

THE DISSIMILATION PRINCIPLE (DP)[2]

Functional expressions taking DNPs as arguments and functional expressions taking CNPs as arguments tend to serialize on the opposite side of their argument expressions.

The SP predicts that VPs, TVPs, PossPhs, and Adpositions (= Prepositions in English) will all occur on the same side of their DNP arguments. Similarly, APs, Relative Clauses, Articles, Determiners, and Quantifiers will tend to occur on the same side of their CNP arguments. And the DP says that these two sides are different. Thus there should only be two characteristic word order types with respect to functional expressions of DNPs and CNPs: Either Function + DNP and CNP + Function or DNP + Function and Function + CNP. Although these predictions are too strong in one respect and so will be weakened slightly below they do make a very substantial number of correct predictions as we shall show and thus do constitute evidence in support of a general correlation between SF and LF.

Thus consider the characteristic word order patterns in Verb Final (SOV) languages, easily the most widespread word order type across different geographical areas and genetic groupings. The relevant word order correlations are:

(52) a. Subject + VP (e.g. John sings; John Mary kissed)
Object + TVP
DNP + Postposition (e.g. the garden-in)
DNP + PossPh (e.g. John's father)
b. Adjective + CNP (e.g. tall man)
Relative Clause + CNP (e.g. the apple eating man)
Article + CNP (the man; this man)
Quantifier + CNP (every man)
Numeral + CNP (two man)

Clearly all the functions on DNPs occur on the same side of the DNP, the right, and all the functions on CNPs occur to the same side, the left. So the SP is fully supported. And these two sides are different, so the DP is

fully supported, at least as far as our current knowledge of the word order correlates goes. Specifically, we know more about some of the word order patterns than we do about others. In particular, as indicated earlier, the internal construction of numeral expressions is complex, nonuniform, and not well known across languages. Further, in all cases here, and below, it should be recognized that we need large scale studies of the word order correlates to make our claims more accurate, and the few such studies we have (see e.g. Hawkins 1979a and 1979b) will surely show that the internal analysis of word order patterning is more complex than indicated above. Nonetheless, relative to the current state of our ignorance, the above mentioned correlates are (with the serious hedge regarding numerals) the best gross statement we have for the structures referred to. (Other correlates, such as the position of Auxiliaries and Main Verbs are not considered here.)

With these qualifications then, consider the second word order type predicted by the SP and DP. It would be:

(53) a. VP + Subject
TVP + Object
Preposition + DNP
PossPh + DNP (e.g. the father of John)
b. CNP + Adjective
CNP + Relative Clause
CNP + Determiner (= Article, Quantifier, or Numeral)

The first two claims in (53a) predict that a language of this type will have VOS as a basic word order. In fact, such languages are rare, although about 12 such cases are known. See Keenan (1978) for a study of their typological properties. On the other hand, though the VOS type is rare, the predictions re word order above are correct (modulo the hedges given above for SOV languages).

But what about the word order types, SVO and VSO not predicted by the two principles? Not only do they exist, but they are both more widespread than the VOS type, though very significantly less so than the SOV type. (54) below summarizes the known word order types characterized in terms of the relative order (where fixed) of Subject, Object, and Verb:

(54) SOV > SVO > VSO > VOS > OVS

The chart below gives a rough breakdown of the distribution of word order types across the major genetic-areal grouping of languages. A major heading is understood to mark a single genetic group unless indicated otherwise. Subheadings mark genetically coherent subgroups.

GENETIC-AREAL DISTRIBUTION OF WORD ORDER TYPES

1. *Australian*
SOV with much freedom of word order

2. *Indo-Pacific* (6 genetically independent groups)
SOV

3. *Austro-Tai*
SVO and VSO both prominant orders. (Some SOV in New Guinea)

4. *Austro-Asiatic*
4.1 *Munda*--SOV
4.2 *Mon-Khmer*--SVO

5. *Sino-Tibetan*
SOV

6. *Dravidian*
SOV

7. *'Boreo-Oriental'* (possibly four genetically independent subgroups)
7.1 *Altaic* (Turkic, etc.)--SOV
7.2 *Uralic* (Samoyedic and Finno-Ugric)--dominantly SOV with some SVO
7.3 *Japanese*, *Okinawan*, *Korean*--SOV
7.4 *Paleo-Siberian*--dominantly SOV

8. *Indo-European* (extant groups)
8.1 *Baltic*--SVO
8.2 *Slavic*--SVO
8.3 *Germanic*--SVO/SOV
8.4 *Italic*--SVO with hints of VSO
8.5 *Indo-Iranian*
8.5.1 *Iranian*--SOV
8.5.2 *Indic*--SOV
8.6 *Dardic*--SOV
8.7 *Greek*--SVO
8.8 *Albanian*--SVO
8.9 *Armenian*--(modern) SOV
8.10 *Celtic*--VSO

9. *Caucasian* (perhaps 2 genetic groups)
SOV dominant, scattered SVO

10. *African* (4 genetically independent subgroups)
10.1 *Khoisan*--SOV
10.2 *Nilo-Saharan*--VSO, SVO, and SOV all attested
10.3 *Niger-Kordofanian*--SVO
10.4 *Afro-Asiatic*
10.4.1 *Semitic* (modern)--SVO
10.4.2 *Berber*--VSO
10.4.3 *Chadic*--SVO
10.4.4 *Cushitic*--SOV

11. *Amerindian* (13 genetically independent subgroups)
SOV probably the dominant order (e.g. Eskimo-Aleut, Na-Dene (Athapascan), Hokan, Aztec-Tanoan, Andean-Equatorial, Macro-Chibchan)
VSO heavily attested: (Salish, Wakashan, Penutian, Oto-Manguean)
SVO is only rarely attested
OVS is known in one case (Hixkaryana, a Carib language in Brazil, see Derbyshire (1977) for discussion)

It is clear from the table that SOV is clearly the dominant order among the world's languages. The relative proportions of SVO vs VSO are hard to evaluate given that I cannot accurately assess just how widespread it is among the Amerindian, though it is clear that it is a dominant order in the genetic groups indicated above. So VSO occurs as a dominant order in several Amerindian phyla, it is one of the dominant orders in Austro-Tai (Polynesian, Philippines, Formosan), and occurs as the dominant order in genetic subgroups of Indo-European (Celtic), Nilo-Saharan (Nilotic), and Afro-Asiatic (Berber). As well it was likely the historical order of older Semitic languages: Classical Arabic, Biblical Hebrew, Babylonian Aramaic. SVO on the other hand, is the dominant order in Europe (ignoring Celtic) though in the phylum as a whole, SOV is a very significant order (Armenian, Dardic, Iranian, Indic). In addition SOV is the dominant order in the extinct IE languages (Hittite) and is the most likely reconstructable order for Proto-IE. SVO is the dominant order in Niger-Kordofanian, a very large group, and occurs as a dominant order in many Austro-Tai language groups (Tai, Indonesian, Javanese). It is the dominant order in Chadic and modern Semitic and occurs sporadically in groups that are otherwise dominantly SOV. I consider it more widespread then than VSO.

VOS only occurs rarely, and then mainly in Phyla which independently present Verb Initial as a major order--Austro-Tai and Amerindian. OVS as noted is cited for only one case, though others related to it may have the same order. We shall not further consider this case here.

What about the word order patterns in SVO and VSO languages other than those involving the subject? Essentially they pattern like we predict for the VOS languages, though more 'doubling' (Hawkins 1979b) occurs, especially in SVO languages (e.g., such languages may have both prenominal and postnominal possessors, as English; both prenominal and postnominal adjectives, as French, etc.). Further and large scale study would be needed to justify that articles and demonstratives occur significantly more frequently postnominally than prenominally, as they are generally considered to do (e.g. Indonesian, Yoruba, Batak, etc.). These observations suggest the following condition:

SUBJECTS FRONT (SF)

The Subject occurs to the left of other major constituents of Ss.

Thus SF is maximally satisfied in both SOV and SVO languages, less satisfied in VSO languages, and still less in VOS and OVS languages. Note further that the SF is a much more specific principle than either the SP or the DP, as the latter quantify over classes of functional expressions, whereas the SF only mentions one specific argument expression--the Subject. If we take then the possible word order types for human language to be those obtained from the set defined by the general principles SP and DP, namely SOV and VOS, and modified in the direction of satisfying Subject Front we obtain (pending of course a more rigorous formulation) a surprising number of correct preditions:

(i) The possible types are SOV, VOS, VSO, and SVO. Just the well attested cases, with the exception of the one OVS language cited.

Note what a strong prediction (i) in fact is. There are 9 function-argument structures considered and thus in principle $2^9 = 512$ possible word order types. Of which only four are predicted to occur, and they do.

(ii) The principles predict that the SVO, VSO, and VOS languages will evidence the same serialization in their function-argument structures with the exception of the placement of the Subject. But all the others should be the same, and this is correct.

(iii) The combined effect of the three principles would appear to correctly predict the observed genetic-areal distribution of the word order types.

Thus the rigorous SOV type satisfies all the principles and is in fact the most widely distributed word order type. SVO fully satisfies Subject Front and fails the Serialization Principle (and the Dissimilation Principle) at just one point: the Subject should follow the VP, that is the TVP + Object. VSO also fails the SP (and thus the DP) at this point, and in addition satisfies Subject Front to a lesser extent than the SVO type and so is less widely distributed. Independently VSO yields a less good correspondence with LF since the TVP + Object does not form a constituent in SF. Finally VOS (and OVS) maximally fail Subject Front and might then be expected to have the least wide distribution.

Despite the logical looseness of prediction (iii) above it does seem to me that the approach we have taken to the word order correlates more successfully relates them to function-argument structure than earlier attempts. Vennemann (1972) is perhaps the first (but see note 2) attempt to systematically correlate word order types with function-argument structure. His approach however is not based on an independently justified logic. He takes for example DNPs as functions with VPs and Prepositions as arguments, and their outputs are supposed to somehow have the same category as the argument. But no natural logic treats, for example, *in* and *in the garden* as having the same type of denotation. Further, Vennemann's approach must rule out the Subject-VP relation in principle, so predictions like (iii) above, however loose, cannot in principle be stated in his framework.

More recently Gil (1979) relates function-argument structures to word order types in the framework of a prosodic theory. His approach assumes our SP and has an analogue of the DP, though much more elaborate. The two approaches are sufficiently different to resist quick comparison, but it is noteworthy that his approach does yield several predictions concerning distinctive prosodic structures in SOV and SVO languages, one of which (greater consonant/vowel ratios in the SOV group) he rigorously supports.

To conclude this section, we might wonder whether the Subject Front principle is anything other than an ad hoc device to get the word order types correctly predicted. Is there any independent reason to expect this principle, and is there any independent justification for it? The answer to both questions I feel is yes, although more work than can be presented here would be needed to justify this claim. So briefly:

As regards motivation, the Subject-VP constituent is the major function-argument structure of simple sentences and might well be expected to have

more perceptual or cognitive salience than the function-argument structures embedded within either the Subject or the VP. Thus, recall from the Meaning-Form Dependency Principle that how we interpret VPs may vary with the choice of Subject. Thus if the VP preceeds the Subject, hearers will have to suspend an exact interpretation for it until the Subject is enunciated. For example, if English were a VP first language there would be sentences beginning with *is flat . . . , is strong . . .* , etc., where the exact sense of the VP could not be interpreted until we knew whether the speaker was talking about roads, beers, voices, or whatever. Further, the problem would quite obviously be more serious in a VOS language than in a VSO language, since more material in a VOS language would be present before the Subject was enunciated. To say for example in a VOS language that *nobody loves both his father and his mother*, we would begin with, *loves the father of his and the mother of his . . .* , without being able to identify the referent of *his*. And if the language allowed sentential objects (the most likely position for complements of verbs of saying) we could begin sentences like thinks that hurt himself he . . . before the Subject phrase was enunciated, requiring the hearer to 'hold in temporary memory' *himself* and *he* until their referents can be established by the Subject. (See Keenan 1978 for further discussion of the disadvantages of placing the Subject at the end of the sentence.)

The motivation suggested above for presenting Subjects to the left of the other major constituents in a sentence does in fact suggest some independent motivation for Subject Front. In the first place, in those languages which do use a VOS order, we in fact never find that the Object position can be filled in surface by a full sentence. To say in those languages (e.g. Malagasy, Fijian, Toba Batak, Ineseño Chumash, etc.) things like *John thinks that Fred will win* a variety of alternatives are available, one of which must be used, and all of which present the thinker before the embedded clause. Thus the language (e.g. Malagasy) may sinply use a VSO order here, in violation of the normal word order. Or it may passivize (Malagasy, Toba Batak) yielding an order like *thought by John that will win Fred*. Or it may utilize any of various devices (clefting, topicalization) for presenting the Subject before the VP.

In the second place, all VSO languages and all VOS languages present means of fronting the Subject. Some of these may be clearly marked structures, e.g., a particle may be inserted after the fronted Subject and before the remaining VP, but such options always exist and, in all cases where I have relevant data (Malagasy, Fijian, K'ekchi (Mayan) and Toba Batak) such fronting devices are more commonly used than, for example, the 'backing' devices (e.g. Right Dislocation) in SOV or SVO languages.

In the third place, even in VSO languages, there will be a variety of ways to form complex (intransitive) verb phrases, e.g. by conjoining two independently intransitive verb phrases. The expected order then would be *sing and dance John*. And sometimes this order can be realized (Malagasy, Isthmus Zapotec, Tamazight (Berber)), but in general it seems more common to break open the VP yielding 'dissonant' structures like *sing John and dance*, where we have a surface coordination between a VP and a full Sentence. And if the VP is a TVP plus Object it is even harder to maintain the VP-Subject order, despite the highly dissonant structures that result.

That is, we would expect to get [*kissed Mary and danced*]$_{VP}$ *John*. (And sometimes we do, as in Malagasy.) But the more usual order is [*kissed John Mary*]$_{S}$ *and danced*.

These cases do then provide independent evidence for Subject Front as a constraint in Universal Grammar on the class of possible word order types.

4. *Theoretical Significance of Correlations between SF and LF.* Why, as linguists, should we be interested in establishing a correlation between the properties of SFs and those of LFs? The answer lies in our goal of trying to represent the linguistic competence of speakers of a language. To know, and thus to learn, a language is to know (learn) much more than simply what the class of grammatical SFs are. Even children from the youngest age do not randomly produce approximations to well formed SFs. Rather they use them with logical (and other) effect. That is, they use them to assert--to make true statements (or statements they intend as true). They use them to deny, and they readily make inferences from what others have said. In other words, in practice, speakers use SFs as LFs. It makes sense to say that what someone has said entails something else, or that it contradicts what someone else has said.

If we represent at least the logical properties of an SF by a set of LFs we may say that what someone knows when he knows a language is a set of pairs <s,t> where s is an SF and t is an LF which represents one of the meanings of s. We may represent the entire language as the set of pairs <s,t> such that s is in SF, t is in LF and t represents one of the meanings of s. (So the language will be a very proper subset of SF x LF.)

(To be sure this idealization abstracts away from much else we know when we know a language, such as most everything we might call 'sociolinguistic', but for the moment this abstraction is sufficient for our purposes.)

To know a language on this view then requires that we know what the class of SFs is, what the class of LFs is, and most important that we know which SFs are paired with which LFs. This last claim can be represented by saying that the ideal speaker knows (in the sense of having internalized) a function which associates each SF s with the set T_s of LFs, each of which expresses one of the meanings of s.[3] In fact, one of the explicit goals of current linguistic theory (at least the Revised Extended Standard Theory, EST) is to define such a function, although work specifically on that topic has in fact been rather scant.

Thus, to account for the linguistic competence of the (ideal) native speaker we shall want to characterize the set of SFs, the set of LFs, and the interpreting function which relates the two. And we shall want to use this characterization to explain what it is that people learn when they learn a language and how it is that children, with a limited and imperfect exposure to the language, nonetheless learn to identify a substantial subset of the pairs (s,t) which constitute the language.

Part of the explanation for this latter fact within generative grammar

has been that humans come biologically equipped with a certain syntactic potential. So the child, or anyone else, is not prepared to accept as a possible set of SFs a very large number of the sets of expressions which can be formally defined. We rightly do not expect that somewhere in the unexplored wilds of New Guinea we will run across a language all of whose sentences have prime length. Thus much work in generative grammar has, rightly to my mind, been concerned to constrain the class of possible grammars the learner can potentially accept. These general constraints would then directly reflect our innate syntactic potential.

Analogous claims hold for the class of possible interpreting functions. We cannot believe that the way humans associate meaning with form, or LFs with SFs, is random. Given a set of SFs for a language, not just any function into a (power) set of LFs is a possible interpreting function for that language. Imagine for example a function F which, for every surface form s, maps s onto a set of LFs none of which mentions any of the LFs which F associates with any of the constituent parts of s. So the LFs which F might associate with *John came early and left late* could be just those that a 'right' interpreting function associated with *Not all unicorns have horns*. Mathematically it is easy to construct such nonadmissible interpreting functions.[4]

So to characterize our competence we want to define the class of possible interpreting functions, and to explain (at least in part) the learnability of human languages, we want to define the class of possible human languages, that is the possible sets of pairs (s,t), as narrowly as possible. Clearly the most constrained class of interpreting functions is the set of identity functions. That is, for any language L, the set of LFs for L would be identical to the set of SFs for L. This would directly explain why speakers can use SFs with logical force (assert, deny, infer, etc.).

Unfortunately this identification does not seem possible on any theory of SF ever propounded. The reason is that in all theories, at least some SFs will be semantically ambiguous and thus correspond to more than one LF. Note that the objects on which the entailment relation is defined (LFs) cannot be ambiguous in the sense of having more than one truth value in some state of affairs. If p were such a formula then we could argue both that p entails p and that p does not entail p, which is to say that the entailment relation itself is not well defined. (To see this, note that given the existence of a state of affairs in which p is true on one reading but false on another it follows that p does not entail p. On the other hand, 'for any state of affairs M, if p is true in M then p is true in M' is itself trivially true. Which is to say p entials p.)

But how badly must the identification of SF with LF fail? Which is to say, how ambiguous are SFs? The question has no obvious answer since it depends at lease on what SFs our theory defines, and as we have seen, that is a complicated matter. Still, one kind of SF which is semantically ambiguous in all theories are those containing semantically ambiguous lexical items of the same grammatical category and subcategory. Thus *John is a bachelor* we are told is four ways ambiguous according as John is a male seal, a knight's helper, etc. And since it can be true on one reading and

false on another, we will need distinct LFs to correspond to each of these readings. Notice however that each of those LFs could be structurally isomorphic to the SF. They might just differ in that within the logic would be four distinct logical common nouns, $bachelor_1$, $bachelor_2$, etc. This type of ambiguity then, while serious, still permits a very strong form of correspondence between SF and LF. Namely, identity up to lexical disambiguation.

But of course linguists and philosophers have been much concerned with nonlexical ambiguities. Many of these however will be eliminated in a reasonably rich theory of SF. Thus most theories would assign *flying planes can be dangerous* two distinct surface forms. (On one, *flying* would be the head of the Subject, and on the other it would be a modifier. On one the head would be marked + plural, and on the other not, etc.) And it appears that many of the classical structural ambiguities are in fact not ambiguous at a level of SF. Note that even a subcategory difference in a lexical item yields a different labelled node in SF. Thus very possibly *the shooting of the hunters* can be naturally assigned to SFs according as the underlying verb *shoot* is transitive or a derived intransitive.

It is worth noting in this regard that elaborating the base, as by proliferating the number of categories and subcategories, does not seriously complicates the grammar as a whole. That is, the class of context free or even context sensitive languages is a very constrained class. If all of English could be given in a context free (sensitive) form the learnability problem would be solved. On the other hand, the class of languages generable by context free bases plus length decreasing filters is enormously larger. Thus to constrain the class of grammars we have an interest in enriching the base and limiting as much as possible the filters used.

On the other hand, there are several sorts of 'possible' ambiguity which have been studied by philosophers and linguists alike which are much less likely to be represented by dintinct SFs. The major cases known to me are the following: (1) cross-reference ambiguity, as in *John told Bill that he was clever*; (2) Quantifier scope ambiguity, as (supposedly) in *Every man loves a woman*; and (3) transparency/opacity mabiguity as in *Suzie wants to marry a Swede* or *John believes that the author of Waverly was the author of Waverly*.

Just how far from isomorphism these cases will force the interpreting function is unclear at the moment. A serious part of the reason is that we have no really clear pretheoretical intuition of what shall count as semantic ambiguity.[5] Thus the position that the *he* in *John told Bill that he was clever* is simply freely generated and its interpretation depends on the context of utterance is at least defensible. So on that view the sentence would not be logically ambiguous. Its 'pragmatic' potential would have to be explicated within a theory of pragmatics or speech usage. However, while many are working in this area, there is not yet, to my knowledge, a sufficiently developed body of science here to be called a theory, and the claim that our sentence is only pragmatically one way rather than logically one way is only a possibility. Similarly one might want to sweep quantifier scope ambiguity under the rug of pragmatics, and reduce the transparency/opacity ambiguity to one of scope of variable

binding operators (an approach which has been reasonably successful for some of the classical opacity cases; but many unsolved problems remain), thus reducing it in effect to the quantifier scope cases.

So it appears thinkable that there could be, at least in principle, something like an isomorphism between an SF and any of its LFs. However, it seems more likely that the mismatch between SF and LF will be greater than that. The ambiguity issue only stated one in principle condition that required a mismatch. But just because a SF is semantically unambiguous it will not follow that, say, its constituent structure corresponds point for point to that of its LF. And further, on some approaches to SF, many surface forms would be generated which would correspond to the null set of LFs. Examples might be *the man was hitting themselves*, *John loves each other*, etc. So on this view the interpreting function itself would in effect act as a kind of filter on the SFs.

So it seems to me then that the best assessment of the relation between SF and LF lies somewhere strictly in between isomorphism and randomness. How close to one pole or the other is an open question, but an important one, since our answer to the problem will determine how successfully we can characterize our linguistic (semantic + syntactic) competence. We should, it seems to me, be looking for theories of SF and LF in which the correspondence is as close as possible, while still accomodating the ambiguity 'facts'. For the closer the correspondence the more we have accounted for a speaker's competence in asserting, inferring, denying, etc. Further, a close correspondence theory gives us perhaps an additional means of explaining how children learn quickly with limited exposure. It not only limits the ways the child may attempt to assign meaning to form but it gives an additional reason for the child's quick learning. Namely, we may assume the child is motivated both to express his needs (and in fact his understanding of the world) and also to understand what others say. That is, there is an intense motivation to be able to perform and comprehend 'semantic' acts. So if SFs are directly useful in coding and in decoding meaning the child is motivated (or at least pushed) into learning the semantically significant properties of SFs. In more concrete terms, the child is motivated to learn the grammatical distinctions between common nouns and intransitive verbs if these distinctions are useful in expressing meanings and understanding the meanings others communicate. If on the other hand, that distinction had no semantic value the child would be less motivated to learn it, and so presumably slower in acquiring it. The lazy child might relapse into conjugating nouns or putting determiners on verbs. We know in fact from psychology (though the cases studied are not directly comparable to the case we are considering) that people learn meaningful material more quickly than meaningless material. To the extent that SFs can be directly interpreted, that is, treated as LFs, the distinctive properties of SFs are meaningful and can be expected to be learned more quickly than if these properties were meaningless.

5. *Conclusion.* Unfortunately, it seems to me, much of the post-*Aspects* work in generative grammar has, in different ways in different cases, moved away from a conception of SF in which it is closely related to LF.

The original conception of syntactic theory in *Syntactic Structures* however asemantic in conception, lent itself very nicely to a close relationship between LF and syntactic form generally. The kernal sentences on that view represented a finite amount of information a language learner had to learn by brute force. The unbounded set of derived structures however could naturally have their meanings represented as a function of what they were derived from since the generating functions (the transformations) were in a naive sense semantically interpretable. Thus a child could learn the meaning of Yes-No Question Formation (one of the generating functions) by learning how the meaning of a small sample of questioned kernal sentences differed in meaning from the kernal sentences they were derived from. Then the meaning of any yes-no question can be ascertained relative to the declarative it is formed from simply by recognizing that it is, syntactically, a yes-no question. Analogous claims held for most of the major transformations in *Syntactic Structures*. A proper semantic theory at this time would have 'simply' faced the task of indicating how the output of a transformation varied in meaning as a function of its input. And this is the sort of thing that logical theories do. Logical derivations by conjunction, disjunction, quantification, and negation do not preserve meaning. Rather we state the meaning of the output as a function of that of the input. Needless to say of course the semantic operations of English would have been at a higher level of complexity than in Classical Logic, but not in principle different.

But as is well known linguistic theory did not develop in this way. By 1964 and 1965 assigning meaning representations to SFs was a recognized goal of the theory, but the theory was reorganized in such a way that only base structures and lexical items were to be semantically interpreted. Taking off from the point, Generative Semantics attempted to argue that the putout of the Base differed very significantly from SFs, and that in fact the base structures (modulo a few differences as regards quantified expressions) resembled the LFs of the classically given first order logic (CL)! This is a terribly negative view of the relation between SF and LF, although that was perhaps not so apparent at the time.

But consider that the entire syntax of LFs in CL can be given in a single paragraph. Here it is: CL has a denumerable set of variables, x_1, x_2, etc. each having the category of Name. It has a set of function symbols each of unique degree $n \geq 0$, and a rule which says that the concatenation of an n-place function symbol with n occurrences of names is a name. The language also has a set of relation symbols each of unique degree $n \geq 1$, and a rule which says that the concatenation of an n-place relation symbol with n names is a formula. And it has two rules for deriving complex formulas. One says that if S and T are formulas then (not S), (S and T), and (S or T) are formulas. And the other says that if S is a formula and x a variable then (for every x, S) and (for some x, S) are formulas. And that's all.

Now the best generative grammars of English we have (Stockwell, Schachter, Partee 1973) run into hundreds of pages, and they are clearly inadequate. Assuming that the meanings expressible in English can be done in CL or a slight extension thereof, we are saying that almost all of English syntax is needless as regards the expression of (logically

representable) meaning. That is, we could have done with a one-paragraph grammar rather than the massive thing we apparently require. And since the diversity of structure present in SFs vastly exceeds the primitive syntax of CL, this view forces a massive correspondence failure between SF and LF. I find this a discouraging view, a kind of modern Babel theory in which Man is punished by God by having to use an absurdly complicated grammar when he could have done with a sleek one-paragraph one.

Nor does this view conform to my experience as one who has been concerned to represent the logical properties of natural language. The more I have looked at natural language the more I have found that almost everything in surface is relevant to the logical interpretation of SFs.

One might expect on the other hand that there would be a more optimistic view of the relation between SF and LF within EST syntax. It is clearly a goal of that theory to map SFs onto LFs, and the closer the correspondence between the two the easier it will be to define the mapping. But I am uneasy about the relation between SF and LF within this theory. On the one hand, in 'Questions of Form and Interpretation' Chomsky (1974: 16) espouses the view that there is a close relation: '. . . the thesis of "absolute autonomy formal grammar" . . . would not imply that there are no systematic connections between form and meaning. No one, I am sure, has ever doubted that there are highly systematic connections and that a major problem of linguistic theory is to determine them.' On the other hand, Chomsky and Lasnik (1977:429) assert that they believe, 'There is . . . some empirical support for the belief that the syntax of LF is close to that of standard forms of predicate calculus . . .' (though that assumption does not appear to me to play an important role in the article).

Given our earlier discussion that an assignment of LFs to natural language expressions is nonobvious in the same sense in which assignment of SFs is, it is certainly unwarranted to assume that LFs have one or another particular shape. To my knowledge no specific attempt has been made within EST syntax to define a function mapping SFs onto descriptively adequate sets of LFs drawn from Classical Logic. We cannot then assess the feasability of defining such a function nor the demands it would make on the language learner.

I can think of only two reasons why, as generative grammarians, we might want to assume prior to an investigation that the LFs of CL are an appropriate range for an interpreting function for natural language SFs. I shall argue here that these reasons are in fact only apparent.

The first is that according to a certain well defined metric of logical complexity (the degree of recursive unsolvability of the predicate logically true) CL is reasonably simple, though not absurdly so. Thus in CL logically true is a recursively enumerable predicate, whereas in second or higher order logics this is not the case. Thus for example in second order logic there can in principle be no complete syntactic characterization of the entailment relation. So if the child (or whoever) only has to know things of first order difficulty, he has to know less than if he has to know things of second or higher order difficulty. This point is surely correct. To claim that the natural logic for English has the full

expressive power of second order logic is to make a stronger claim about our logical competence than to claim that first order logic is sufficient.

However, the order of a logic is not tied to any particular syntactic instantiation, such as that in CL. The order of a logic is strictly determined by the set theoretical type (relative to the universe of discourse) of the variables in the logic. What makes first order logic first order is that we can only quantify over members of the universe of discourse U. In second order logic we can quantify over subsets of U, and so we can say things like 'For every object x there is a property P such that P(x)' and so on. So a logic which only allows quantification over U is first order, and that commits us to very little concerning the class of LFs the logic defines. It only requires that we have variables of a certain type but otherwise we are free to construct the rules defining the LFs as we please. For example, the extensional logic proposed in *Logical Types* is in all essential respects a first order logic. (The range of the variables in that logic appears to be sets of higher types, but in fact they are so constrained as to be in a one-one correspondence with the members of U.)

Secondly, the question of close correspondence between SF and LF is in principle independent of the question of autonomy of syntax. I take the autonomy thesis to be essentially that our syntactic capabilities are innately determined up to some level of specificity and that these capabilities are independent of other things, such as the semantic ends we make them serve. In fact construed in this way the autonomy thesis seems to me clearly true. There is to be sure a serious empirical matter as to how much of say English syntax is innately determined and how much represents a 'free' choice within the bounds of our innately determined possibilities. But the in principle question seems clear: only humans use language, so what more natural explanation than that humans differ biologically in this respect from other mammals?

But from the autonomy thesis alone nothing <u>follows</u> concerning the relation between SF and LF. It certainly does not follow for example that agreement pairs do not correspond to function-argument structures in the way we have claimed in this paper. Nor will it follow that certain function-argument structures are not commonly expressed by SFs which exhibit regular left-right order relations among their parts. Nor will it follow that the interpretation of complex SFs is completely unrelated to the interpretation of their constituent parts.

Whether correspondences exist between SFs and their LFs is an empirical matter. To say that there is a massive correspondence failure is to say that man has made very poor use of his syntactic endowment. It is to say that man makes very many syntactic distinctions which are in fact irrelevant to the expression of meaning; and that he fails to code in his syntax distinctions which are semantically relevant. I can believe that we have not made perfectly efficient use of the syntactic instrument we come endowed with; that idioms for example do present meaningful parts whose meaning is (synchronically) irrelevant to the meaning of the entire idiom. And I can believe that there may be certain systematic ambiguities in language, such as the opacity/transparency ones cited earlier, which are not regularly coded in SF.

But I cannot believe that the syntactic elaboration present in each language can be collapsed into the pitiful syntax of CL. How will we represent the voicing systems of the Philippines? The tense systems of Bantu? The switch reference systems of New Guinea and the Americas? Where shall we put the deixis systems present in all languages? Where will imperatives and hortatives go? And subjunctives? And gerunds? What about the rules which most if not all languages have which convert Ss and VPs into NPs and Adjectives? And Adjectives into abstract nouns? Where to put the mass nouns, and noun compounds? And do we really not logically distinguish common nouns from intransitive verbs? And adjectives from adverbs?

Surely if we have all this syntax in our heads we will use it for some semantic effect. It may to be sure be difficult to explicate just what the semantic effect is in particular cases, but rather than accept a punishment theory--all that syntax is useless--I would recall the German proverb often cited by Einstein: 'The Lord is subtle but he is not mean'.

NOTES

[1]Several alternatives for the analysis of PPs are in fact suggested in *Logical Types*. The version fully presented there was one in which Prepositions combined with either intransitive verbs (VPs) or transitive verbs (TVPs) to extend the class of DNP arguments they take. The analysis in which this semantics is presented with Prepositions as functions taking DNPs as arguments has been worked out by Faltz (Dept. of Linguistics, UCLA). For simple DNPs that analysis is equivalent to the one in *Logical Types*. If the DNPs are properly quantified there may be relative scope differences between the DNP object of the Preposition and the Direct Object of transitive verbs.

[2]It is interesting to note that perhaps the first scholar to investigate word order correlates, Abel Bergaigne, in 1875, arrives at a very similar analysis: they can be expressed by a 'double formule: le terme qualifiant precede le terme qualifie et le terme regi precede le terme regissant'. The quote is cited in Holland (1976:413).

[3]We need not require that T_s be the entire set of LFs each of which expresses one of the meanings of s. It is sufficient that T_s be a set of LFs such that any logical form which represents a meaning of s is logically equivalent to one in T_s. T_s then can be expected to be finite for any s.
Note incidentally that whether the set of pairs $\langle s, T_s \rangle$ is recursive is completely independent of whether an arbitrary chosen logical form is valid, that is, true in all interpretations of the logic.

[4]For example, let I be a correct interpreting function for a set of SFs. So for each SF s, I(s) is the set of descriptively adequate LFs for s. Enumerate the set of SFs (so for each SF we may associate a unique natural number n). Pick any wierd numerical function, say the function f which maps each natural number n onto $100 + n^3$. Define a 'wrong'

interpreting function I* which maps each surface form s onto I(f(n)), where n is the number in the enumeration of s, and I(f(n) is the set of LFs which I associates with the SF whose number is f(n).

[5]For example, as regards entailment we do have a good pretheoretical intuition. Namely, an (unambiguous) sentence S entails an unambiguous sentence T iff T is true in all the cases in which S is true. Model theoretic semantics is an attempt to formalize this notion by formally representing the notion 'true in a state of affairs'.

The best pretheoretical intuition behind the notion 'ambiguity' which I can come up with is at best only a necessary condition for a sentence to be ambiguous. Namely, if S is ambiguous between meanings A and B (however represented) then a speaker on a literal, sincere, etc., use clearly intends either A or B but not both. Thus the speaker cannot respond 'I don't know' to the query 'Do you mean A, or B?' We may attempt to formally represent this intuition by designing an interpreting function for SFs which maps the pretheoretically judged ambiguous ones onto a set containing non-logically equivalent LFs. But the descriptive adequacy of the mapping is only as good as is our pretheoretical intuition of ambiguity, and that in fact is not very good.

REFERENCES

Anderson, S. R. 1971. On the role of deep structure in semantic interpretation. Foundations of Language 7. 387-96.

------. 1977. On mechanisms by which languages become ergative. In Li 1977, 317-63.

Capell, A. 1971. Arosi grammar. Pacific Linguistics Series B--No. 20. Canberra: The Australian National University.

Chomsky, N. 1974. Questions of form and interpretation. Manuscript. Department of Linguistics, MIT.

------ and H. Lasnik. 1977. Filters and control. Linguistic Inquiry 8. 425-505.

Comrie, B. 1973. The ergative: Variations on a theme. Lingua 32. 239-53.

------. 1978. Ergativity. In Lehmann 1978.

------ and S. Thompson. 1978. Lexical nominalizations. Manuscript. Department of Linguistics, UCLA.

Derbyshire, D. C. 1977. Word order universals and the existence of OVS languages. Linguistic Inquiry 8. 590-99.

Dixon, R. M. W. 1972. The Dyirbal language of North Queensland. Cambridge: Cambridge University Press.

Gary, J. and E. L. Keenan. 1977. On collapsing grammatical relations in universal grammar. Syntax and semantics: Grammatical relations, Vol. 8, ed. by P. Cole and J. Sadock, 63-121. New York: Academic Press.

Gil, D. 1979. Prosodic structure and linguistic universals. Manuscript. Department of Linguistics, Tel Aviv University.

Hawkins, J. 1979a. Implicational universals as predicters of word order change. Language 55.

------. 1979b. Two types of word order universals: Implicational and

distributional. Manuscript. Department of Linguistics, University of Southern California.
Holland, G. 1976. The shift from postposition to preposition: Evidence from Early Greek. Proceedings of the Second Annual Meeting of the Berkeley Linguistics Society.
Keenan, E. L. 1976. The logical diversity of natural languages. Annals of the New York Academy of Sciences 280. 73-92.
------. 1978. The syntax of subject-final languages. In Lehmann 1978.
------ and L. Faltz. 1978. Logical types for natural language. Occasional Papers in Linguistics. No. 3. Department of Linguistics, UCLA.
Kimenyi, A. 1976. A relational grammar of Kinyarwanda. Ph.D. dissertation. Department of Linguistics, UCLA.
Lehmann, W. (ed.). 1978. Syntactic typology. Austin: University of Texas Press.
Li, C. (ed.). 1975. Mechanisms of syntactic change. Austin: University of Texas Press.
------ and S. Thompson. 1975. Subject and word order in Wappo. Manuscript. Department of Linguistics, UCLA.
Montague, R. 1973. The proper treatment of quantification in ordinary English. Approaches to natural language: Proceedings of the 1970 Stanford Workshop on Grammar and Semantics, ed. by J. Hintikka, J. Moravcsik, and P. Suppes, 221-242. Dordrecht: Reidel.
Moravcsik, E. 1978. Ergative and accusative patterns. Manuscript. Department of Linguistics, University of Wisconsin.
Munro, P. 1976. On the form of negative sentences in Kawaiisu. Proceedings of the Second Annual Meeting of the Berkeley Linguistics Society, 308-19.
Murane, E. 1974. Daga grammar. Norman, Okla.: Summer Institute of Linguistics.
Partee, B. 1976. Montague grammar. New York: Academic Press.
Payne, J. 1978. Phrasal conjunction. Manuscript. The Typology Project, Department of Linguistics, UCLA.
Reinhart, T. 1976. On understanding poetic metaphor. Poetics 5. 383-402.
Silverstein, M. 1976. Hierarchy of features and ergativity. Grammatical categories in Australian languages, ed. by R. M. W. Dixon, 112-71. Canberra: Australian Institute of Aboriginal Studies.
Stockwell, R., P. Schachter, and B. Partee. 1973. The major syntactic structures of English. New York: Holt, Rinehart, and Winston.
Timberlake, A. 1975. The nominative object in Finnish. Lingua 35. 201-30.
------. 1977. Reanalysis and actualization in syntactic change. In Li 1977, 141-77.
Vennemann, T. 1972. Analogy in generative grammar: The origin of word order. Paper read at the XIth International Congress of Linguists. Bologna, Italy.

LANGUAGE UNIVERSALS IN LINGUISTIC ARGUMENTATION

James D. McCawley
University of Chicago

1. *Introduction.* I will begin this not very comprehensive survey of the ways in which language universals have figured and/or ought to have figured in linguistic argumentation by outlining certain assumptions and prejudices that will color what I say below.

First, I take the position that all language universals are implicational universals. Only under that assumption can one avoid arbitrary decisions as to what will count as a language. Most work on language universals has systematically ignored child language, written language, sign language, nonnative language (pidgins, Roman Jakobson's English, etc.), and the language of aphasics and schizophrenics. For example, linguists happily accept the claim that all languages have nasal consonants[1] even though they are perfectly aware of languages that do not, for example, American Sign Language. It is of course reasonable for linguists to ignore ASL here: the universal relates to the vocal medium of spoken languages. However, the irrelevance of ASL to some language universals does not make it irrelevant to all discussion of language universals--it clearly is relevant to discussions of universals of constituent order or of semantic distinctions in the lexicon. If 'language' is understood broadly and universals are taken to be implicational, exclusion of any types of language from the domains of universals must be for cause: the antecedent of the implicational universal must give the grounds on which the particular varieties of language are to be excluded (e.g., 'If the medium of expression of a language is vocal, it will have nasal consonants, excludes ASL by virtue of its medium of expression being nonvocal), and the linguist must attempt to state universals in their greatest generality, i.e., by specifying as narrowly as possible the conditions that would remove a variety of language from the applicability of the putative universal.

Second, I wish to dissociate myself from an assumption that is so popular among linguists that it is difficult to find anyone who disputes it, namely the assumption that people who talk the same have the same linguistic competence. I have recently been advocating (McCawley 1976a; 1977) a conception of language acquisition in which many details of acquisition are random or are influenced by ephemeral details of one's linguistic experience. Such a scheme for language acquisition need not lead to gross inhomogeneity in a linguistic community, since there is ample opportunity for revision of learning that has made the speaker grossly divergent from his neighbors. Moreover, it is easy to think of alternative linguistic rules and underlying forms that yield exactly the same well-formedness data and exactly the same pairings of meaning and expression; indeed, linguists are perpetually arguing about such

alternative analyses, e.g., the analysis in which the English regular plural ending is /iz/ and a rule deletes its vowel under one set of circumstances, versus the analysis in which the ending is /z/ and a rule inserts a vowel under other circumstances. Maybe some people have learned plurals the one way and other people the other way. The assumption that all normal adult members of a linguistic community have the same internalized analysis in such cases is gratuitous. In the rare cases where linguists have looked for interpersonal variation in language, they have generally found it. For example, Haber (1975) reports that speakers who ostensible form English plurals the same way give a broad range of responses on tasks requiring the formation of plurals of novel words, with each speaker having his own ways of dealing with novel plurals. There is also considerable individual variation in the morphemic relations that speakers of English perceive, as one can readily verify by asking one's friends whether *pulley* is related to *pull* or *tinsel* to *tin*.

Third, and closely related to the last point, I claim conscientious objector status in the ongoing war against 'excessive power' of grammatical devices. We have all been taught to limit our descriptive devices as tightly as possible, preferably to those that were good enough for our scientific forefathers, since seemingly innocuous devices may turn out to harbor within them the dread Turing machine. It has accordingly become common for linguists to attempt to resolve disputes among competing analyses by drafting sweeping restrictions on grammars so as to give one of the competing analyses a legal monopoly. I will argue below that many proposed language universals have served only to allow linguists to construct cheap arguments for their favorite analyses and that those arguments have given the illusion of significance only because their alleged role in the war effort against 'excessive power' has obscured important respects in which they are extremely implausible.

2. *Universals as self-serving legislation.* I will begin this largely negative survey of the use of language universals in linguistic argumentation by criticizing and recanting a couple of particularly bad arguments that I have offered in earlier works. In McCawley (1973a), I presented a number of arguments for treating vowel height in terms of a single 3 or 4 valued feature rather than in terms of two binary features such as Chomsky and Halle's (1968) [+ high] and [+ low]. The bulk of my arguments were based on phenomena in which constant increments of height played a role, e.g., the raising of the first element of a long vowel by one degree of height in eastern Finnish dialects and the lowering of the second element of a long vowel by one degree in certain Lappish dialects. I then introduced my final argument by saying 'What would give a really crushing case for the nonbinary proposal would be to show some major restriction on rules which could be imposed under the nonbinary proposal but not under the binary proposal' and proceeded to discuss a supposed example of such a restriction: a requirement that phonological rules have a unitary 'structural change', i.e., a restriction excluding rules that specify changes in more than one feature specification at a time. Under the binary proposal, that universal would be violated by rules such as the Finnish and Lappish rules just discussed, whereas a nonbinary alternative involved a unitary structural

change:

(1) Eastern Finnish 'breaking'

a. binary version

$$\begin{bmatrix} -\text{ high} \\ \alpha\text{ low} \\ \beta\text{ front} \\ \partial\text{ round} \end{bmatrix} \quad \begin{bmatrix} -\text{ high} \\ \alpha\text{ low} \\ \beta\text{ front} \\ \partial\text{ round} \end{bmatrix}$$

$$1 \rightarrow \begin{bmatrix} -\text{ low} \\ -\alpha\text{ high} \end{bmatrix}$$

b. nonbinary version

$$\begin{bmatrix} \alpha<3\text{ high} \\ \beta\quad\text{ front} \\ \partial\quad\text{ round} \end{bmatrix} \quad \begin{bmatrix} \alpha\text{ high} \\ \beta\text{ front} \\ \partial\text{ round} \end{bmatrix}$$

$$1 \rightarrow [\alpha{+}1\text{ high}]$$

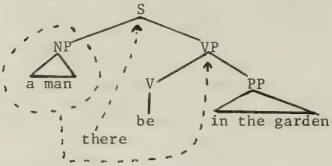
I also argued for reanalyses of certain apparent instances of rules involving changes of two or more feature specifications. Specifically, I argued that in cases such as the laxing, deaspiration, and despirantization of syllable-final consonants in Korean, the effects of the rule were predictable side effects of a unitary change (in this case, the rule basically just made all syllable-final consonants unreleased, and the other changes represented predictable concomitants of unreleased articulation) and showed how cases such as compensatory lengthening could be given a similar treatment, namely by taking the structural change not to be a change in the feature composition of segments but a change in the timing of phonetic events, in this case, postponement of the onset of a following consonant, with the changes of feature composition of the segments involved being predictable consequences of that single change of timing.

While I was thus able to present a fairly good case that one could impose a universal restriction that a phonological rule have a unitary structural change, I did not take up the question of whether that restriction had anything else to recommend it beyond its apparent viability as a universal. First, what that universal implies depends heavily on how one answers an extremely difficult question that linguists have said little about and are not in a position to say anything very coherent about, namely that of how one can justify a particular individuation of rules: how one can justify saying that a particular formula represents a single entry in the grammar rather than being merely a misleading way of representing several such entries (or perhaps, pieces of several such entries). To the extent that rules with multipart structural changes can be simulated by sets of rules with unitary structural changes (which is a large extent unless quite stringent additional constraints are imposed on what is allowed as a possible rule), the proposed universal is in effect merely a policy to divide among several rules the structural change of what one might otherwise have considered a single rule. Second, if the structural change of a rule is to be related in any even indirect way to something that a speaker does or may do in the course of speech production, it is implausible to suppose that doing two such things together should be ruled out; if anything, structural changes that always go together ought to be kept together in the internalized grammar for maximum economy in the use of the grammar in production and comprehension.

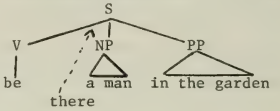
A second place in which the putative universal that rules have unitary structural changes played a role in my argumentation is the

paper that I published under the title 'English as a VSO language' (McCawley 1970).[3] I pointed out there that a number of transformations require an additional step if they apply to SVO structures that is unnecessary if they apply to VSO structures.[4] For example, if *there*-insertion applies to an SVO structure, it must both move the underlying subject into a postverbal position and insert *there* in subject position, whereas if it applies to a VSO structure, it need only insert a *there* before the subject (the underlying subject will remain in postverbal position in surface structure, since only the first of the verb's sisters is moved into preverbal position by the rule that yields the surface SVO order):

(2) SVO *there*-insertion



VSO *there*-insertion

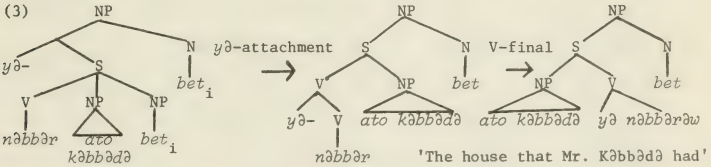


I ought to have argued against the SVO formulation of *there*-insertion on the grounds that it misses a generalization: all rules that create a new subject also demote the old subject to nonsubject status, and the SVO formulation of *there*-insertion makes the demotion of the old subject seem like an added complication of the phenomenon rather than a predictable side effect of the other part of the structural change.[5] What I did instead was argue for the VSO formulation on the grounds that it allowed one to enforce a universal restriction that transformations must have unitary structural changes. Besides sharing the implausibility of its phonological analogue, discussed above, this putative universal has the further defect of committing one to an analysis of the passive that has nothing to recommend it. As I pointed out in 'English as a VSO language', the SVO version of passive requires an extra part to its structural change that is unnecessary in the VSO version: not only must it move the underlying subject into the *by*-phrase but it must also move the underlying postverbal NP into subject position. However, if Passive is to have a unitary structural change, the markers of the passive (passive *be* and the *by*) must not be inserted by the Passive transformation but must be present as such in the input to Passive. Underlying structures in which those elements are present have in fact been proposed widely (an underlying *by* is proposed in Chomsky 1965 and assumed in Chomsky's subsequent works; an underlying *be* is adopted in Hasegawa 1968, Lakoff 1971, and, I am ashamed to say, McCawley 1970). However, the arguments that have been given for such underlying structures rest on preposterous premises, particularly the assumption that transformations cannot 'build structure'. Thus, for example, the argument that Chomsky (1965) seemed to find the weightiest for having passive *by* present in deep structure is that, given his policy on where node labels come from (namely that they must be the labels of corresponding deep structure nodes), that is the only way he can get the *by*-phrase to be a prepositional phrase or even a syntactic constituent. What little substance these arguments might

have vanishes if one adopts Anderson's (1976) position that principles of syntactic category assignment relate not just to deep structure but to all stages of a derivation (e.g., anything consisting of a preposition and a NP is a PP, regardless of what it is derived from) and the further position that transformations apply so as to achieve conformity with the language's surface target configurations.[6]

Another highly dubious universal that has been pressed into service in arguments for linguistic analyses that allow that universal to be enforced is the putative universal that no quantifiers need be used in formulating rules. For example, in Bach's (1970) otherwise superbly argued case for deriving the surface SOV word order of Amharic from an underlying VSO order, the following argument appears. In Amharic, restrictive relative clauses precede the head noun and are marked by a prefix *yə-* on the final verb of the relative clause. Bach maintains (and let us grant him these assumptions, since it is a different point of the argument that I wish to attack) that, regardless of whether Amharic has underlying VSO, SVO, or SOV word order, in deep structure the relative clause must precede the noun as it does in surface structure (a reasonable assumption) and *yə-* must precede the relative clause (a not so reasonable assumption).[7] Bach notes that VSO underlying order allows one to formulate the *yə-* attachment rule without the use of quantifiers (its structural description is simply [X *yə* V Y], whereas with SOV word order the structural description of *yə-* attachment must make use of a quantifier to specify that the material after the verb to which *yə-* is attached contains no verb, i.e., that *yə-* is attached to the last verb (the last verb in SOV order being, of course, the first verb in VSO order). (3) gives a sample derivation according to Bach's preferred analysis:

(3)



'The house that Mr. Kəbbədə had'

A restriction excluding rules whose formulations include quantifiers would of course significantly narrow the class of possible grammars. However, it would narrow the class to such an extent as to make language impoverished in comparison with other cognitive domains, such as visual perception. A condition such as 'where X contains no V' in a rule with a structural description [... X V Y] or [Y V X ...] amounts to an instruction to match the symbol V to the first V or the last V in the given domain. The sketches of perceptual mechanisms given in Miller and Johnson-Laird (1976) involve wide reliance on searches through structured domains, especially searches for the first or last element that meets a given condition. While I am in no position here to propose any specific integration of psychological mechanisms in perception and in language use, I hold that it is implausible to suggest that organizational features that

figure widely in perception and in nonlinguistic knowledge are systematically excluded from language.

3. *Universals and underlying constituent order.* A much more promising use of language universals in argumentation, one also illustrated in Bach's Amharic paper, is the use of implicational universals and tendencies as the basis for assigning to languages that appear to violate the universals an underlying order different from their surface order. For example, Bach makes much of the atypical status of Amharic among surface SOV languages, e.g., that it has more propositions that postpositions and that it has forwards Gapping. Amharic can be made less of an anomaly by treating it as having VSO word order for the bulk of its syntax (and thus acting like a VSO language with regard to most of its syntactic rules), with only relatively superficial syntactic rules presupposing an SOV order that is derived from VSO order by a transformation moving each verb to the end of its clause.

I have one serious objection to such invocations of language universals in establishing deep constituent orders, namely that they involve a gratuitous assumption that there is a deep constituent order. I attribute the prevalence of that gratuitous assumption to two principal causes: first, a hasty and mistaken judgement that unordered structures are outlandish abominations (this judgement may in many instances be a symptom of DENDROPHOBIA, a malady that is manifested in a tendency to think in terms of strings rather than of trees whenever possible, and at times even when not possible; Chomsky appears to suffer from chronic congenital dendrophobia), and second, a mistaken application of Postal's (1968) 'naturalness principle', according to which underlying linguistic structures have the same formal nature as superficial structures. Postal's principle provides a rationale for taking underlying forms of morphemes to be composed of segments and for taking those segments to be composed of specifications of values for features that also occur in surface phonology, rather than, say, having underlying segments be indivisible units or involve 'abstract' features that have no phonetic correlates. However, the principle only excludes units and structural characteristics that do not figure in superficial linguistic structure--it does not require that all the units and structural characteristics found in superficial structures also occur at deeper levels of linguistic structure. Thus, for example, the principle does not rule out underlying forms in which some or all phonological segments are unspecified with regard to voicing or length. In McCawley (1976b), I have argued that constituents of underlying syntactic structures must be allowed to be unspecified for morphological characteristics (such as person, number, case, or definiteness) that they must have determinate specifications for in surface structure. For example, I argued that underlying structures must be allowed in English in which nouns are indeterminate with regard to number, in view of the fact that while there is as much reason to derive (4a) by conjunction reduction from conjoined NP's as there is to employ conjunction reduction in anything else, corresponding NP's in which conjunction reduction has not applied must be determinate as to whether each of the two composers wrote one or more than one quartet:

(4) a. The quartets of Eierkopf and Misthaufen

b. The quartet(s) of Eierkopf and the quartet(s) of Misthaufen

By all applicable tests for ambiguity, (4a) is unspecified rather than ambiguous as to whether each composer wrote more than one quartet,[8] and thus the nouns in the structure to which conjunction reduction applies to derive (4a) must be unspecified for number, though if conjunction reduction does not apply there must be a number specification on each noun. Postal's naturalness condition, as I understand it, would not exclude an underlying structure for (4a) in which both conjuncts are unspecified for number. Nor would it exclude analyses involving underlying structures that are unspecified for anything else for which surface structures are determinate, such as, in particular, constituent order.

Thus, a fair proportion of arguments that have been offered for underlying word orders different from surface word order are at best arguments that if the language has an underlying word order at all, it is the one that the argument claims to give support to. Accordingly, there is a serious question as to what claims about underlying word or order can be justified without reliance on the gratuitous assumption that there is one, and the concomitant question of whether proposed implicational universals about word order can be given interpretations that will make them applicable to underlying structures that in some cases are wholly or partially unordered. I will make no attempt to answer the latter question here, though I commend it to the attention of my fellow linguists. On the former question, the best I can do at the moment is to cite two reasonably clear cases, one of an argument for an underlying word order different from surface order that is independent of the assumption that there is an underlying word order, the other of a well-known argument for a specific underlying word order for which an alternative analysis is available that requires no linear order of constituents at any level other than surface structure.

Of Bach's arguments for underlying VSO order in Amharic, the following are the strongest, in that it is difficult to attribute the distrubution of the morphemes in question to anything other than the underlying constituent order that Bach proposes. Bach argues that the *yə-* relative clause marker and the *yə-* genitive marker are the same morpheme, on the basis of a shared morphological idiosyncracy, namely that both delete when preceded by a preposition. While *yə-* is prefixed to the last verb of a relative clause, it is prefixed to the first word of a genitive NP:

(5) *yə-yohannis bet* 'John's house'
[*yə-tillik̩ šum*] *bet* 'the house of a big chief'
GEN-big chief house

This is predicted by Bach's proposed underlying word order as long as his hypothesized relative clause reduction in genitive NP's (i.e., deletion of 'which . . . has', e.g., 'which John has' → 'John's') applies before *yə-*attachment.[9] After deletion of the hypothesized 'have' and before movement of the verb to the end of the relative clause, the first word of the relative clause is the verb in the case of a full relative clause, but is the first word of the genitive NP if

the relative clause has been reduced to a genitive NP. Furthermore, no factor other than word order appears to be shared by the words to which *yə-* is attached; for example there is no generalization in terms of the notion 'head', since while the verb can most plausibly be regarded as the head of the full relative clause, the first word of a multiword genitive NP is never the head of either the genitive NP or the relative clause from which the genitive NP is derived. The facts are exactly parallel for the definite accusative marker *-n*, which is suffixed to the first word of a NP that begins with anything other than a (full) relative clause but is suffixed to the last word (the verb) of a relative clause that begins the NP:

(6) *liju-n* 'the child'
tilliku-n konjo lij 'the big beautiful child'
tillikinna konjo yəhonəwi-n lij 'the child who is big and beautiful'

Under Bach's proposed analysis, the word to which *-n* is suffixed is the word with which the NP begins after relative clause reduction but before movement of the V to the end. Again, no alternative factor such as the head/adjunct distinction provides any apparent basis for a generalization as to what word gets the suffix.

On the other hand, the considerations that Koster (1974, 1975) gives as grounds for analyzing Dutch as having SOV deep constituent order can be reconciled with an analysis that agrees in essentials with Koster's but makes no reference to constituent order in anything but surface structure. In Dutch, adverbs and prepositional phrases following the verb in a main clause can often occur in either order (7a), though when they precede the verb in a subordinate clause they can occur in only one of the two orders (7b):

(7) a. *Jan dacht tijdens de pauze aan zijn vader.*
Jan dacht aan zijn vader tijdens de pauze.
'Jan thought of his father during the intermission'
b. *Piet zei, dat Jan tijdens de pauze aan zijn vader dacht.*
**Piet zei, dat Jan aan zijn vader tijdens de pauze dacht.*
'Piet said that Jan thought of his father during the intermission'

Koster explains this odd fact on the basis of the possibility of these elements following the verb in subordinate clauses. Specifically, either or both may follow the verb, but when both do they are in the opposite order to that which they exhibit when they precede the verb:

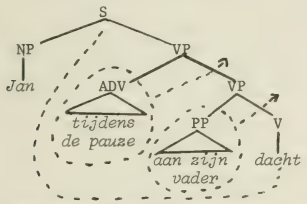
(8) *Piet zei, dat Jan aan zijn vader dacht tijdens de pauze.*
Piet zei, dat Jan tijdens de pauze dacht aan zijn vader.
Piet zei, dat Jan dacht aan zijn vader tijdens de pauze.
**Piet zei, dat Jan dacht tijdens de pauze aan zijn vader.*

With SOV deep constituent order, the facts about main clauses can be seen as parallelling those about subordinate clauses: the second sentence in (7a) arises not through permutation of the prepositional phrases but through movement of one of them to the right of the (basically final) verb, as in the first sentence of (8). Movement of the verb into second

position then leaves the two prepositional phrases adjacent but in the opposite of their original order.

However, the generalization that the word order possibilities in main clauses reflect the possibility of putting material after the verb in subordinate clauses can be incorporated into an analysis that involves unordered constituents in all derivational stages but surface structure. Suppose that, as in de Haan (1976) and in distinction to Koster, we treat adverbial elements as sisters of their respective VP's and treat the occurrence of *tidjens de pauze* after the verb and the occurrence of *aan zijn vader* after the verb as resulting from two different processes: in the one case an analog to English adverb placement, allowing an adverb to either precede or follow its sister VP, and in the other case an analog to English heavy-NP-shift, allowing a sufficiently 'heavy' daughter of a VP node to appear at the end of that VP. In terms of Koster's underlying constituent order, the derivational possibilities could then be summed up in the following diagram:

(9)



The movement of the V into second position in main clauses involves a change in constituent structure. This is particularly clear from consideration of sentences in which something other than the subject appears in first position, e.g., *Tijdens de pauze dacht Jan aan zijn vader*, in which *dacht Jan aan zijn vader* presumably is not a VP and thus the V must be something other than a daughter of a VP node. Note then that Dutch word order can be described in the following terms, which involve no reference to underlying word order: (i) in a main clause, one NP or adverbial expression is made a daughter of the S node; (ii) in a main clause, the finite verb is made a daughter of the S node; (iii) if a VP node directly dominates a V, the V is last within the VP, except that a heavy daughter of the VP node may optionally follow the V, (iv) if a VP node directly dominates a VP and an adverb, they may occur in either order, and (v) if a S node directly dominates a V, a VP, and something else, then the V is second and the VP third.[10] This alternative analysis is in the spirit of Koster's, in that the structure of subordinate clauses is more basic than that of main clauses, but it takes the difference between them as basically one of constituent structure rather than of word order.

4. *Language universals and the notion 'explanation'.* Koster's analysis provides an EXPLANATION of the facts that prepositional phrases following the verb in a Dutch main clause can occur in more orders than are possible when they precede the verb in a subordinate clause. An

explanation purports to show why things are the way they are, i.e., to show why some range of alternative states of affairs are not manifested. Any purported explanation of linguistic phenomena thus carries with it implications to the effect that certain types of language are impossible, e.g., Koster's explanation of the Dutch facts implies the impossibility of a language that allows the same possibilities as Dutch for word order in subordinate clauses, but only allows adverbs in main clauses to occur in an order that they could have if they preceded the verb in a subordinate clause. Thus, the status of a putative linguistic explanation should be regarded as shaky unless an appropriate cross-linguistic survey has been done, in which languages sharing the factors allegedly responsible for the given phenomena are investigated to determine whether they also exhibit those phenomena.

Regrettably few supposed explanations of linguistic phenomena have been accompanied by appropriate cross-linguistic data. One well-known 'explanatory' device that comes to grief when one tries to test it cross-linguistically is rule schemata such as Chomsky's (1970) $\overline{X}$ -> Spec_X $\overline{X}$, which collapses several phrase-structure rules in a single formula (NP -> Det $\overline{N}$, 'Pred P' -> Aux VP, AP -> Degree $\overline{A}$). The claim that language is acquired in terms of such rule schemata supposedly explains the fact that determiners, auxiliaries, and degree expressions in English all occur to the left of their heads. If the notation (as part of a language-acquisition faculty) provides a real explanation of that fact, then a cross-linguistic survey should reveal a tendency for determiners, auxiliaries, and degree expressions all to occur on the same side of their heads (i.e., either all before the head or all after the head). There is in fact no such correlation. While there are many languages in which determiners precede nouns and auxiliary verbs precede main verbs, as in English, there are also many languages where determiners precede nouns and auxiliary verbs follow main verbs (for example, Japanese and Turkish) and many languages where determiners follow nouns and auxiliary verbs precede main verbs (for example, Malay and Swahili), but there are few languages in which determiners follow nouns and auxiliary verbs follow main verbs (Somali is one such language). This suggests that what Chomsky's schema supposedly explains is not really a 'fact' but a conjunction of several unrelated facts that by accident share similarities.

One exemplary instance in which a cross-linguistic survey is used to verify an explanation of language-particular facts is provided by Cole (1974). Cole disputes Postal's (1972) claim that the question (10a) and the relative clause in (10b) are bad for the same reason, namely that an item is moved over something coreferential to it:

(10) a. *Who$_i$ did the claim that he$_i$ is a fraud surprise?
b. *The man who$_i$ the claim that he$_i$ is a fraud surprised has left town.

Cole maintains that (10a) is ungrammatical because its derivation involves a stage in which there is backwards pronominalization with indefinite antecedent and is thus ungrammatical for the same reason as are such examples as (11):

(11) *The claim that he$_i$ was a fraud surprised $\left\{\begin{array}{l}\text{someone}_i\text{.} \\ \text{every member}_i\text{ of the} \\ \text{committee.}\end{array}\right\}$

That cannot be the reason for the ungrammaticality of (10b), since relative pronouns are presumably definite and thus at no stage of the derivation of (10b) is there backwards pronominalization with an indefinite antecedent. Cole proposes to account for the ungrammaticality of (10b) by invocation of a device that linguists have tended to reject out of hand as a cop-out and a failure to capture a generalization, namely a 'rule of analogy' whereby relative clauses like that of (10b) are ungrammatical simply by virtue of their looking like ungrammatical questions. Indeed, if only facts about the variety of English treated by Postal are considered, it is hard to see how any support for Cole's analysis could be found. However, Cole's proposal carries with it cross-linguistic implications that Cole proceeded to investigate and which yielded extremely solid support for his analysis. Specifically, if Cole is correct as to what is responsible for the ungrammaticality of (10b), then in languages that have WH-movement and do not allow backwards pronominalization with indefinite antecedent, sentences like (10a) should always be bad, but sentences like (10b) should be bad only if the language provides a basis for the analogy to which Cole attributes the ungrammaticality of (10b), i.e. analogs to (10b) should be good in languages in which relative clauses do not look like questions (and thus, all such languages should provide counterexamples to Postal's putative explanation of the English facts), whereas they may be bad in languages whose relative clauses do look like questions (though they need not be bad in all such languages--the fact that a language provides the basis for a given analogical rule does not imply that the language must have the rule). And in fact, in languages whose relative clauses do not look like questions, e.g., German and Chinese, relative clauses like that of (10b) are good. Moreover, the split does not go along genetic lines, e.g., English behaves like French and unlike German.

5. *The cleansing power of universality.* The most common responses to embarassing facts have generally been to ignore them and to dispute them. In recent years a third response has in certain linguistic circles come to rival those two responses in popularity, namely the response of claiming that the embarassing facts are universal. This ploy, which might be summed up in the slogan 'Make the dirty stuff universal' is illustrated by the following passages:

(12) 'Notice that if the filter in question belongs to U[niversal] G[rammar], then its proper formulation is irrelevant to the theory of filters; that is, it will have no bearing on the question of the proper format for presenting filters and restrictions on possible filters'. (Chomsky and Lasnik 1977: 451) 'Notice, incidentally, that this conclusion [that the Specified Subject Condition requires a formulation in terms of the semantic notion of 'Agent' rather than the syntactic notion of 'Subject'], if correct, would not affect the hypothesis that transformations do not refer to semantic relations but only to bracketing of phrase markers (see the opening

> discussion) even if the semantic notion of 'agency' plays a role in determining the applicability of transformation'. (Chomsky 1973:124)

Chomsky (both solo and with Lasnik) is willing to allow the applicability of transformations to depend on nonsyntactic factors and is willing to admit filters of grossly different types from those countenanced by the 'theory of filters' that he promises to develop, just as long as the otherwise offensive conditions and filters are universal.

I take Chomsky at his word when he suggests elsewhere that language universals are genetically determined. Note, though, that both universal characteristics within a species and individual variation involve genetic determination. Typically, gross anatomical structures are universal but fine details of structure are subject to genetically determined individual variation, e.g., all normal human beings have two arms involving humerus, ulna, and radius, connected to hands having the standard numbers and arrangement of carpals, metacarpals, and phalanges, but on the other hand no two persons have the same fingerprints. Williams (1956) provides a useful survey of details of human anatomy in which there is considerable genetically determined individual variation, e.g., the shape, capacity, and location of the stomach. The moral to be drawn from these considerations is that genetically determined universals provide loci for individual variation, even of genetically determined individual variation, and no realm of structure can be regarded as immune to individual variation in its finer details. I conclude from this that if biology is as deeply implicated in language as Chomsky says it is, then individual variation (and presumably also differences between populations) should be possible in any realm of linguistic structure. If, for example, Chomsky's suggestion is correct that a version of the Specified Subject Condition involving the notion 'agent' is universal and if that universal is genetically determined, then it would be surprising not to find variation among individuals or between populations as to its fine details, e.g., as to how broadly 'agent' must be interpreted.

6. *Benediction.* In this paper, I have suggested at several points that there may be more diversity in language than linguists generally allow for. This does not mean that there are fewer interesting linguistic universals than linguists have generally suggested there are or that there is any conflict between the search for universals and the search for diversity. Indeed, the characteristic of language universals that I regard as most valuable is that they make it possible for the linguist to focus his attention in such a way as to identify specific types of diversity that would otherwise go unnoticed.

NOTES

[1]For expository purposes, I ignore here a couple of languages indigenous to the state of Washington that may in fact be real counterexamples to this claim.

[2]Compare here the fact that a putative universal that all grammars must contain an even number of rules could be enforced simply by adopting a policy of adding a vacuous rule to any grammar that has an odd number of rules.

[3]It could more appropriately have been called 'English as not an SVO language': the bulk of the arguments were at best arguments that certain rules applied to structures in which the subject and the remaining material were not on opposite sides of the verb; thus the arguments were consistent with SOV order. The one argument that was specifically directed towards establishing verb-initial order does not provide a clear case for that conclusion (see McCawley 1973b:226-7 for details).

[4]I could of course have said 'SOV structures' instead; cf. n. 3.

[5]This is in fact the insight that lies at the bottom of the relational grammar treatment of relation-changing transformations; see Perlmutter and Postal (1977).

[6]Under this position, as contrasted with the 'structure preserving' framework of Emonds (1976), there can be gross discrepancies between the allowable deep structure configurations and the allowable surface configurations. My principal criticism of Emonds' approach is that it rests on a gratuitous assumption that language-particular combinatoric restrictions (= 'base rules') relate to deep structures rather than e.g., to surface structures and that it thus forces one to adopt highly specific and otherwise unmotivated deep structures in order to account for phenomena to which, as far as I can see, deep structure is irrelevant. For example, the phenomenon of Affix-hopping is a reflection of the morphological fact that English verbal affixes are suffixes. Emonds' position requires him to set up deep structures containing empty affix positions so that Affix hopping can be structure-preserving, which, by elimination, it must be in his rule typology. (I should, by the way, say 'the phenomena of Affix-hopping', in view of Akmajian and Wasow's (1975) demonstration that *-en* and *-ing* must be attached to verbs by a rule that applies under different conditions and at a different stage of derivations than the rule that attaches tenses to verbs.)

[7]Bach in this argument displays the common tendency of transformational grammarians to place disproportionately heavy emphasis on whether an affix is a prefix or a suffix. I maintain that the status of a bound morpheme as a prefix or suffix is a matter of its morphology rather than of syntax and that syntactic rules should be neutral as to whether morphemes that they introduce or move are prefixes or suffixes: the rule should attach the bound morpheme to the relevant item, with its morphology determining whether it emerges as a left sister of a right sister of the item to which it is attached. Indeed, while Bach underplays the suffix *-n*, to be discussed below, and concentrates instead on the prefix *yə-*,the facts about *-n* provide an even solider argument for underlying verb-first order, since the argument about *-n* does not rest on a possibly vulnerable claim of morphemic identity the way that the argument about *yə-* (as we will see in section 3) does.

Note, incidentally, that a suffix may be positioned in terms of the beginning of a complex constituent with which it is combined. For example, Latin *-que* 'and' is suffixed to the first word of the conjunct in which it appears, as in *senatus populusque romanus* 'the senate and the Roman people'. This fact provides one of several arguments that *-que*, like *et*, basically precedes the conjunct but is adjoined to the immediately adjacent word of that conjunct but is adjoined to the immediately adjacent word of that conjunct and, by virtue of its morphological status as a suffix, appears at the end of that word.

[8]It is of course ambiguous with respect to another distinction, that between individual authorship (cf. *the operas of Verdi and Puccini*) and joint authorship (cf. *the operettas of Gilbert and Sullivan*). This ambiguity is manifested more clearly in the NP *the operas of Verdi and Boito*.

[9]This condition need not involve any extrinsic rule ordering: it is equivalent to the condition that *yə-* must appear on a word that appears overtly in the relative clause.

[10]This sketch is an oversimplification, e.g., I have not taken up word order in questions. The rule of Verb-raising (Evers 1975), which adjoins a nonfinite verb to the next higher verb, requires some complication in this analysis, since the lower verb must become a right sister of the higher verb. This means that if the imposition of a word order is to be postponed until surface structure, the word order rules must have access to information distinguishing the roles of the two sister verbs, e.g., identification of one of them as head and the other as adjunct.

REFERENCES

Akmajian, Adrian, and Tom Wasow. 1975. The constituent structure of VP and AUX and the position of the verb *be*. Linguistic Analysis 1.205-45.

Anderson, Stephen R. 1976. Concerning the notion 'base component of a transformational grammar'. In McCawley 1976c:113-28.

Bach, Emmon. 1970. Is Amharic an SOV language? Journal of Ethiopian Studies 8.9-20.

Chomsky, Noam A. 1965. Apsects of the theory of syntax. Cambridge, Mass.: MIT Press.

------. 1970. Remarks on nominalization. Readings in English transformational grammar, ed. by P. Rosenbaum and R. Jacobs, 184-221. Waltham: Ginn.

------. 1973. Conditions on transformations. A festschrift for Morris Halle, ed. by S. Anderson and P. Kiparsky, 232-86. New York: Holt, Rinehart and Winston. (Page references are to the reprint in N. Chomsky, 1977. Essays on form and interpretation, 81-160, Amsterdam and New York: North Holland.)

------ and Morris Halle. 1968. The sound pattern of English. New York: Harper and Row.

------ and Howard Lasnik. 1977. Filters and control. Linguistic Inquiry

8. 425-504.
Cole, Peter. 1974. Backward pronominalization and analogy. Linguistic Inquiry 5.425-43.
Emonds, Joseph. 1976. A transformational approach to English syntax. New York: Academic Press.
Evers, Arnold. 1975. The transformational cycle in Dutch and German. Ph.D. thesis, University of Utrecht.
de Haan, Ger. 1976. Regelordening en domainformuleringen op transformaties. Lijnen van taaltheoretisch onderzoek, ed. by G. Koefoed and A. Evers, 279-302. Groningen: H. D. Tjeenk Willink.
Haber, Lyn. 1975. The muzzy theory. Papers from the 11th regional meeting, Chicago Linguistic Society. 240-56.
Hasegawa, Kinsuke. 1968. The passive construction in English. Language 44.230-43.
Koster, Jan. 1974. Het werkwoord als spiegelcentrum. Spektator 3.601-18.
------. 1975. Dutch as an SOV language. Linguistic Analysis 1.111-36.
Lakoff, Robin. 1971. Passive resistance. Papers from the 7th regional meeting, Chicago Linguistic Society, 149-62.
McCawley, James D. 1970. English as a VSO language. Language 46.286-99. Also in McCawley 1973b:211-228.
------. 1973a. The role of notation in generative phonology. The formal analysis of natural languages, ed. by M. Gross, M. Halle, and M. Schutzenberger, 51-62. The Hague: Mouton. Also in McCawley 1979: 204-16.
------. 1973b. Grammar and meaning. Tokyo: Taishukan; and New York: Academic Press.
------. 1976a. Some ideas not to live by. Die neueren Sprachen 75. 151-65. Also in McCawley 1979:234-46.
------. 1976b. Morphological indeterminacy in underlying syntactic structure. Proceedings of the 1975 Mid-America Linguistics Conference, 317-26. Lawrence: University of Kansas. Also in McCawley 1979:113-21.
------. 1976c. Notes from the Linguistic underground (Syntax and Semantics 7). New York: Academic Press.
------. 1977. Acquisition models as models of acquisition. Studies in language variation, ed. by R. Fasold and R. Shuy, 51-64. Washington: Georgetown University Press.
------. 1979. Adverbs, vowels, and other objects of wonder. Chicago and London: University of Chicago Press.
Miller, George, and Philip Johnson-Laird. 1976. Language and Perception. Cambridge, Mass.: Belknap Press.
Perlmutter, David M., and Paul M. Postal. 1977. Toward a universal characterization of passivization. Proceedings of the third annual meeting of the Berkeley Linguistics Society, 394-417.
Postal, Paul M. 1968. Aspects of phonological theory. New York: Harper and Row.
------. 1972. A global constraint on pronominalization. Linguistic Inquiry 3. 35-59.
Williams, Roger. 1956. Biochemical individuality. New York: Wiley.

LINGUISTICS IS ABOUT LANGUAGES

Bernard Comrie
University of Southern California

The claim that linguistics is about languages might seem, at first sight, to be a truism, following from the etymology of the word *linguistics*. At the time when I was a graduate student, however, in the late sixties and early seventies, for a student of linguistics to be interested in languages was almost a vice. The then-prevalent paradigm within transformational-generative syntax, though ostensibly interested in language universals, argued that the best way to study language universals, to hypothesize and test claims about language universals, was via the detailed study of a single language. For rather obvious practical and social reasons, this one language was usually English.

The most cogent argumentation in favor of this approach to language universals was given by Noam Chomsky, and the following may serve as a very much abbreviated, but I believe essentially correct, account of this argumentation. Linguists working within this paradigm were not unaware of the apparent conflict between 'language universals' and 'detailed work on a single language'. Chomsky's view of language universals was, and is, intimately connected with his views on innate ideas, i.e. species-specific genetically transmitted abstract structural properties that severely delimit the class of possible languages that human beings can learn naturally. There is thus an isomorphism between language universals and innate ideas. Chomsky claims that these innate ideas are extremely abstract in nature, relating to properties that are not typically manifested directly in the surface structures of sentences. Thus the study of language universals would involve, at least as a first step, the construction of the appropriate set of abstract principles and, since these are not readily observable in surface structure, this must depend on detailed in-depth study of a limited area: broader, but more superficial, investigation would probably fail to uncover the very abstract generalizations in which transformational grammarians were typically interested.

Moreover, it was claimed, these abstract principles are not for the most part expected, i.e. not the sort of restrictions one would expect a priori to find in a communication system. This led to a parallel argument justifying the equation of abstract principles hypothesized on the basis of detailed work on a single language and language universals/innate ideas. If a given structural principle is sufficiently abstract, i.e. sufficiently removed from generalizations that can be made directly about the surface data, then it is unclear, so the claim would go, how the child could induce such generalizations from the data with which he is presented as a first-language learner. In this sense, the abstract generalizations would be unlearnable, therefore, if they exist, they must be innate. Since children are able to learn equally well whatever human language is spoken in their community, these innate ideas must be common to all humans, whence universal.

In this lecture, I wish not to attack directly the bases of this line of argument, but rather to present, in summary form, some recent pieces of research in language universals which, it seems to me, argue strongly for a very different methodology, since they point to claims about language universals that can only be verified or disconfirmed by examining data from a wide range of languages. Several of the pieces of research discussed are my own, although I have also included other research, in particular word order typology. Although most of the pieces of research in question relate to problems rather different from those posed during the heyday of transformational-generative grammar, some of them are very similar to such problems, for instance the discussion of nominalizations in section 3. The discussion of individual points is necessarily very brief in this lecture, although I have included references to more detailed treatments.

1. *Word Order Typology.* Starting with the publication of Joseph H. Greenberg's paper on word order typology (Greenberg 1966), one topic that has been of great interest to linguists working on language universals has been word order typology. In particular, it seemed at one time that of the six logically possible ways of ordering the elements Subject, Object, and Verb of a sentence (as in (1)), certain did not occur in the languages of the world, i.e. it might be possible to establish a language universal restricting the possible orderings of S, O, and V admissible in a human language.

(1) (a) SOV (b) SVO (c) VSO
(d) VOS (e) OVS (f) OSV

Greenberg (1966) noted that almost all human languages have as their basic word order one of the first three types listed in (1), i.e. either SOV (like Japanese and Turkish) or SVO (like English and Swahili) or VSO (like Welsh and Classical Arabic); of these three, VSO is itself considerably rarer than the first two. Greenberg, as always being careful in his formulation of language universals, did not exclude the possibility of types (d)-(f), and indeed cited some languages that had been claimed in the literature to have these word orders. However, it did seem that some kind of generalization was lurking among these data, suggesting a universal of the type: 'in basic word order, subjects always precede objects'. Since the languages cited by Greenberg as possible exceptions to this generalization often turned out, on closer inspection, not to be exceptions --often in the sense that it was simply difficult or impossible to tell what their basic word order is--, this putative universal soon gained widespread acceptance.

Subsequent work on a variety of languages, however, reduced the class of non-occurring basic word-order types. The investigation of Malagasy syntax carried out in greatest detail by Edward L. Keenan suggested that Malagasy has as its basic word order VOS, i.e. type (d) of (1), as in the following example:

(2) *nahita ny mpianatra ny vehivavy.*
saw the student the woman
'the woman saw the student' (not: 'the student saw the woman')

Although some attempts were made to exclude the relevance of Malagasy and other VOS languages, for instance by claiming that the clause-final noun phrase is a topic rather than a subject, these objections were successfully countered by Keenan (1976), which establishes that Malagasy does indubitably have a noun phrase with the range of properties one expects a subject to have, and which indubitably occurs sentence-finally in basic word order.

At this stage, with types (a)-(d) attested, there still remained a possible generalization concerning word order, namely: 'in basic word order, the object is never initial', since this is the single feature separating (a)-(d) from (e)-(f). The question that thus arises is whether or not there exist object-initial languages. Again, only work being carried out on a wide range of languages enabled an answer to be given to this question. Desmond C. Derbyshire, in answer to a polemical statement by Geoffrey K. Pullum to the effect that object-initial languages do not exist, suggested that Hixkaryana, a Carib language of Brazil, is an OVS language, as illustrated in:

(3) *toto yonoye kamara.*
man he-ate-him jaguar
'the jaguar ate the man' (not: 'the man ate the jaguar')

Detailed justification of this claim was presented subsequently in Derbyshire (1977), countering possible arguments that OVS is a nonbasic order in Hixkaryana. Further collaboration between Pullum and Derbyshire, concentrating on languages of the Amazon basin, suggests that the sixth logical possibility may well also be attested, for instance in Apurinã, as in example (4) (Pullum & Derbyshire 1978):

(4) *kema Pedro na- nika.*
tapir Pedro NEG ate
'Pedro didn't eat the tapir' (not: 'the tapir didn't eat Pedro')

What this research on word order typology, covering a range of languages, shows is that a certain putative universal is in fact untenable, the universal being either that subjects must precede objects, or that objects may not be initial. Either of these universals could, incidentally, have been given a ready psychological interpretation, for instance in terms of the primacy of agents over patients in the conceptualization of situations. Although the strong form of the universal proves to be untenable, there is still a weaker version that may be valid, namely that basic word order in languages tends to follow the following principles: (i) subjects tend to occur sentence-initially; (ii) subjects tend to precede objects. As tendencies (statistical universals), these two principles are quite salient.

2. *Ergativity.* One problem that has long intrigued linguists interested in syntactic systems different from those found in the majority of European languages is the problem of ergativity. In its original sense, the term ergative system refers to a system of noun phrase case-marking where one case (absolutive) is used for both intransitive subject and transitive direct object, whereas another case (ergative) is used for the

subject of a transitive verb, as in the following examples from the Australian language Dyirbal (Dixon 1972):

(5) *yaŋa baninju.*
man-ABS came-here
'the man came here'

(6) *djugumbil baninju.*
woman-ABS came-here
'the woman came here'

(7) *djugumbil yaŋa-ŋgu balgan.*
woman-ABS man ERG hit
'the man hit the woman'

The ergative (or, more fully, ergative-absolutive) case-marking system differs from that found in most European languages that have case-marking of noun phrases, where the same case (nominative) is used for both intransitive and transitive subjects, and a different case (accusative) for direct objects. Even in English, where little case-marking of noun phrases remains, this system (nominative-accusative, or simply: accusative) remains for pronouns:

(8) *he* (NOM) *came here.*

(9) *she* (NOM) *came here.*

(10) *he* (NOM) *hit her* (ACC).

Although the existence of ergativity as a morphological phenomenon has been known since the late nineteenth century, its implications for the rest of the syntactic structure of languages with this morphological system has only recently begun to be understood. In earlier work, quite far-reaching hypotheses were built upon the morphological facts, such as the claim that the notion of subject is quite irrelevant to ergative languages, or the claim that speakers of ergative languages have a quite different way of thinking from speakers of accusative languages (or even that the former are incapable of thinking properly at all)! When detailed work was begun on the syntax of languages with morphological ergativity, mainly from the middle of the twentieth century, it seemed at first that the syntax of these languages did not reflect the morphological opposition absolutive/ergative, but rather followed a nominative/accusative (subject/direct object) opposition as in nominative-accusative languages. Thus in such languages as Basque, Georgian, and Hindi, a number of syntactic processes refer to the subject of a sentence, and the reference of the term subject of the sentence is essentially the same as in English or Latin, although in ergative languages this goes against the morphology of noun phrases. Thus ergativity appeared to be a purely morphological quirk of certain languages.

The publication of R. M. W. Dixon's grammar of Dyirbal, however, shattered this complacent view (Dixon 1972). This was the first detailed syntactic account to be published of an Australian aboriginal language,

and from this account it emerged that, where English (and Basque) have a range of rules referring to subjects (transitive and intransitive), Dyirbal has a range of rules referring to absolutives (i.e. intransitive subjects and direct objects, but not transitive subjects), i.e. that in Dyirbal the syntax mirrors the ergative morphology. As an illustrative example, I shall use coordination reduction. In both English and Dyirbal, if two clauses contain an identical noun phrase, it is sometimes possible to delete the occurrence of this noun phrase in the second clause (in English, joining with a conjunction, e.g. *and*; in Dyirbal, usually by simple juxtaposition). Both languages have a syntactic condition on this deletion, but, as will be seen from examples (11)-(12), the condition is different in the two languages:

(11) *the man hit the woman and* (sc. *he/*she*) *came here.*

(12) *djugumbil yaɻa-ŋgu balgan, baninju.*
woman-ABS man ERG hit came-here
'the man hit the woman and she/*he came here'

In English, if one conjoins a transitive sentence and an intransitive sentence, then the subject of the intransitive sentence may be deleted if it is coreferential with the subject of the preceding transitive sentence; thus the omitted subject of the second clause of (11) must be interpreted as *the man*, not as *the woman*. In Dyirbal, on the other hand, if one conjoins a transitive sentence and an intransitive sentence, then the intransitive subject of the second clause may be deleted if it is coreferential with the object of the first clause, i.e. with the noun phrase in the absolutive case. Thus in example (12), the omitted subject of the second clause must be interpreted as *the woman*, and cannot be interpreted as *the man*, any more than the English sentence *the man hit the woman and came here* can be given the interpretation 'the man hit the woman and she came here'.

On the basis of this general difference between English and Dyirbal syntax, Dixon at first hypothesized that the syntax of a given language would be either entirely on a nominative-accusative basis (as in English) or entirely on an ergative-absolutive basis (as in Dyirbal), with no possibility of combining the two systems within a single language. However, further work on ergative languages--including work by Dixon himself on other Australian languages--showed this strict dichotomy to be untenable. This may be illustrated by some examples from Yidinj, another Australian language, which happens to be a geographic neighbor of Dyirbal, although it is very different from Dyirbal in a number of respects (Dixon 1977). Case-marking of noun phrases in Yidinj is on an ergative-absolutive basis, as can be seen in the following examples:

(13) *waɋu:dja maŋga:nj*
man-ABS laughed
'the man laughed'

(14) *bunja maŋga:nj.*
woman-ABS laughed
'the woman laughed'

(15) *bunja:ŋ wagu:dja wuṛa:nj.*
woman-ERG man-ABS slapped
'the woman slapped the man'

There are, however, some exceptions to this generalization. In particular, personal pronouns have rather a nominative-accusative case-marking system:

(16) *ŋaju maŋga:nj.*
I-NOM laughed
'I laughed'

(17) *njuntu maŋga:nj.*
you-NOM laughed
'you laughed'

(18) *ŋaju njuninj wuṛa:nj.*
I-NOM you-ACC slapped
'I slapped you'

(In Dyirbal too, pronouns have a nominative-accusative case-marking system, but syntactically they behave exactly like other noun phrases, i.e. according to the ergative-absolutive system.) As in English and Dyirbal, there is a syntactic restriction on deletion of noun phrases in coordinate constructions in Yidinj:

(19) *bunja:ŋ wagu:dja wuṛa:nj, maŋga:nj.*
woman-ERG man-ABS slapped laughed
'the woman slapped the man, and he laughed'

(20) *ŋaju njuninj wuṛa:nj, maŋga:nj.*
I-NOM you-ACC slapped laughed
'I slapped you, and I laughed'

Examples (19) and (20) differ in that the former has nonpronominal noun phrases (ergative case-marking), while the latter has pronouns (accusative case-marking). What is interesting, however, is that the syntactic possibilities for deletion exactly mirror the case-marking system: in (19), the omitted intransitive subject is interpreted as coreferential with the absolutive direct object of the first clause (as in Dyirbal), whereas in (20) the omitted intransitive subject is interpreted as coreferential with the nominative subject of the first clause (as in English). In other words, Yidinj combines aspects of both accusative syntax and ergative syntax within the one language (and indeed within the one construction--coordination reduction). Thus a strict dichotomy between languages with ergative syntax and languages with accusative syntax is untenable. Indeed, one finds that even in a radically ergative language like Dyirbal there are traces of accusativity (case-marking of pronouns; deletion of the addressee in imperatives, where deletion may be of the second person subject, intransitive or transitive), while radically accusative languages like English have traces of ergativity (such as the suffix *-ee*, forming nouns that refer to intransitive subjects, e.g. *escapee* 'one who has escaped' or objects of transitive verbs, e.g. *employee* 'one whom someone else employs').

The work described in this section on ergativity in a range of languages might seem to have been essentially negative in terms of establishing language universals, since at each stage a stronger hypothesis has had to be replaced by a weaker one: the first hypothesis was that all languages are syntactically nominative-accusative; the second that each language is either completely nominative-accusative or completely ergative-absolutive in syntax; the third that a given language can combine aspects of accusativity with aspects of ergativity, i.e. essentially the null-hypothesis with regard to ergativity. However, more recent work on how aspects of accusativity are combined with aspects of ergativity suggests that, once one recognizes the existence of 'split ergativity' of this kind, there are several generalizations that can be made concerning the possibilities for combining ergativity and accusativity. For instance, if the case-marking system of a language is split by the pronominal/nonpronominal distinction, then it is always the pronouns that have accusative case-marking while other noun phrases have ergative case-marking. If some morphological or syntactic feature is split by tense or aspect, then it is always the case that the past/perfective/perfect tense/aspect correlates with ergativity while the nonpast/imperfective/nonperfect tense/aspect correlates with accusativity. These and other generalizations, which significantly reduce the range of possible human languages, are discussed in greater detail in, for instance, Comrie (1978b).

3. *Nominalizations*. From the very inception of transformational-generative grammar, one construction that has attracted widespread attention was the action nominal, as in English *the enemy's rapid destruction of the village*, because of the apparent discrepancy between surface syntactic structure and deep syntactic structure. In surface structure, such constructions are clearly noun phrases: the head constituent *destruction* is a noun, and its attributes are genitives (in English, prenominal with *'s* or postnominal with *of*) and adjectives. More abstractly, however, such action nominals must be related to sentences like *the enemy rapidly destroyed the village*, despite the differences in surface structure: in the sentence, we have a verb with arguments in the nominative and accusative (in English, distinguished only for some pronouns), and qualified by an adverb (e.g. *rapidly*) rather than an adjective (e.g. *rapid*). Early transformational work concentrated on the parallels between action nominals and sentences, deriving both from the same, essentially sentential, deep structure. It was Noam Chomsky himself who suggested abandoning this transformational analysis for action nominals, pointing out that previous work had neglected the fact that the internal syntactic structure of an action nominal exactly parallels that of other noun phrases, a fact ignored, or treated as accidental, by the transformational approach to this construction (Chomsky 1970). Under Chomsky's revised analysis, action nominals would have throughout their derivation the internal syntactic structure of noun phrases, and there would be no question of transforming sentential syntactic structure into nominal syntactic structure.

A corollary of this revision is that there should be a rigid distinction between sentential syntax and nominal syntax, since Chomsky's proposal rests on the assumption that action nominals should either have the internal structure of noun phrases in toto, throughout the derivation,

or (as is indeed the case in some languages) that they should have the syntax of sentences in toto, throughout the derivation. In a more wide-ranging study of action nominals in a variety of languages, I came to the conclusion that, despite some initial plausibility, this consequence of Chomsky's analysis is not tenable, which in turn calls into question the motivation for his analysis in English (Comrie 1976).

If Chomsky's analysis were correct, then one would expect that where languages differ in the internal structure of noun phrases, they would differ in the same way in the internal structure of action nominals, even where they do not differ in basic sentence structure. There are indeed examples of this situation. In English, the sentences and action nominals (21)-(24) are all possible; note that none of the nominals violates the general restriction that a head noun may have up to one prenominal genitive attribute (usually in *'s*) and up to one postnominal genitive attribute (with *of*)--both may also occur simultaneously:

(21) *the enemy destroyed the village.*

(22) *the enemy's destruction of the village*

(23) *the village was destroyed by the enemy.*

(24) *the village's destruction by the enemy*

In Russian, the basic sentence structure is the same, at least in basic word order, with the addition that subject and direct object are usually distinguished by the nominative/accusative case distinction:

(25) *vrag razrušil derevnju.*
enemy-NOM destroyed village-ACC
'the enemy destroyed the village'

(26) *derevnja byla razrušena vragom.*
village-NOM was destroyed enemy-INSTR
'the village was destroyed by the enemy'

With action nominals, Russian has in common with English the general rule that both 'subject' and 'direct object' of an action nominal appear in the genitive, e.g. *prijezd pojezda* 'the arrival of the train' (*pojezda* is genitive of *pojezd* 'train'), *razrušenije derevni* 'the destruction of the village'. However, Russian does not, in general, share with English the possibility of having distinct prenominal and postnominal genitive attributes: Russian has only one genitive position, which is usually postnominal. Chomsky's account of action nominals would thus predict that Russian should have no literal translation of English example (22) (*the enemy's destruction of the village*), although it should have a literal translation of (24) (*the village's destruction by the enemy*), involving only one genitive. This prediction is in fact borne out:

(27) **razrušenije derevni vraga*
destruction village-GEN enemy-GEN

(28) *razrušenije derevni vragom*
destruction village-GEN enemy-INSTR

(Example (27) is grammatical in the sense 'the destruction of the enemy's village', with stacked genitives, though this is irrelevant to the present issue.)

Unfortunately for Chomsky's approach, there are other languages where the parallelism between noun phrase and action nominal syntax does not go through. One such language is Turkish. In Turkish sentences, the subject appears in the nominative, the direct object (at least, if definite) in the accusative, and the verb appears sentence-finally, agreeing with the subject:

(29) *Hasan mektuplar-ı yaz -dı*
Hasan-NOM letters ACC write PAST-3SG
'Hasan wrote the letters.'

In noun phrases which are not action nominals, there can be one genitive attribute, which stands before the head noun; the head noun takes a possessive suffix agreeing with the genitive:

(30) *Hasan-ın baba -sı*
Hasan GEN father his
'Hasan's father'

Turning now to the internal structure of action nominals in Turkish, we observe that they have one feature in common with other noun phrases: the 'subject' noun phrase stands in the genitive, not the nominative, and the head noun takes the appropriate possessive suffix. However, the 'direct object' of the action nominal stands in the accusative, just as if it were an argument of a verb:

(31) *Hasan-ın mektuplar-ı yaz -ma -sı*
Hasan GEN letters ACC write ACTNOM his
'Hasan's writing of the letters'

In Turkish, then, the action nominal combines features of sentential syntax with features of nominal syntax, and there is no rigid distinction between the two. In passing, we may note that the English construction often called the gerundive nominal evinces a similar structure, at least in the written language, since its subject takes the genitive ending *'s*, whereas its object remains prepositionless:

(32) John's reading the book (surprised us all).

As with the discussion of ergativity in section 2, it might seem that the conclusion to this section is essentially negative, having served to destroy a universal rather than to establish one. However, the account of action nominals given in Comrie (1976), of which this section is a summary, still leaves for the most part intact the generalization that action nominals can combine, in different proportions in different languages, sentential syntax with nominal syntax, rather than having

some radically different third kind of syntactic structure. So this account, which is required by the data in preference to Chomsky's, still significantly restricts the range of possible syntaxes for action nominals vis-à-vis the null-hypothesis.

4. *Definite Direct Objects.* One criticism that is often made of work looking at language universals on the basis of data from a wide range of languages is that, in achieving breadth of coverage, accuracy in the description of the individual languages is sometimes sacrificed, i.e. analyses are proposed for a given language on the basis of a cursory examination of the facts for that language, where more in-depth analysis of the language might well lead to rejection of that analysis. This is, of course, a methodological point, rather than one of principle: ideally, one should have in-depth studies of a wide range of languages. Since the individual researcher or even research team disposes of only a finite amount of time, research on universals in practice requires a weighting between depth and breadth of investigation, since the perfect combination of full depth and full breadth is unattainable in practice. One of the main claims of this lecture is that the pendulum should swing more in the direction of breadth of coverage.

Moreover, there are some instances where a broad coverage, of several languages, can suggest solutions to problems in the description of individual languages that might otherwise have eluded the linguist looking only at one language in depth. I shall illustrate this with a problem involving so-called definite direct objects; for further details, see Comrie (1978a). In a number of languages, of which Turkish is one, there is a special morphological case used for direct objects, but only if the direct object in question is definite; otherwise, direct objects stand in the same morphological form as subjects, i.e. in the citation form. Thus with the noun phrase *öküz* 'ox' in Turkish we have the following contrast:

(33) *Hasan bir öküz aldı.*
Hasan an ox bought
'Hasan bought an ox'

(34) *Hasan öküz-ü aldı.*
Hasan ox DEFDO bought
'Hasan bought the ox'

In (33), the direct object is indefinite (note the indefinite article *bir* in Turkish, *an* in the English translation), so there is no 'definite direct object' suffix. In (34), the direct object is definite (note the definite article in the English translation), so we have the definite direct object ending *-ı* (with vowel harmony variants *-i*, *-u*, *-ü*; in certain environments, this vowel is preceded by *y* or *n*).

In general, those direct objects which are case-marked in this way in Turkish correspond to our intuition of what constitutes a definite direct object, to the usual general linguistic definition of definiteness. However, there are some exceptions, such as the following, where the direct object is clearly not definite in the usual sense:

(35) *Hasan hangisi-ni aldı?*
Hasan which DEFDO bought
'which one did Hasan buy?'

(36) *Hasan bir öküz-ü aldı.*
Hasan an ox DEFDO bought
'Hasan bought an ox'

Traditional grammars of Turkish, in particular those written for English-speakers, are well aware of examples of this kind, and the solution they usually adopt is to say that in Turkish 'definiteness' means something different from what it means in English. If one were to take this line to its logical conclusion, then we would be defining technical terms ad hoc for each individual language; although we might use the same term (such as definite) in speaking of two languages, the two uses of the term would be effectively mere homophones, since 'definite' in one language might have nothing in common with 'definite' in any other. This would obviously put an end to any study of language universals, since terms would have no meaning outside the discussion of an individual language, but its implications would be even more far-reaching: under this approach, there could be no substantive theory of general linguistics. The overall negative effect of this assumption would thus be very far-reaching.

An alternative approach, and one that turns out to be promising both for the general theory of definiteness and for the uniform analysis of the Turkish examples, would be to try and isolate similarities between definiteness in its strict sense and the so-called definite direct object in Turkish, these similarities then being what control the identity of morphological form of the direct object in Turkish examples (34)-(36). Following the approach to definiteness proposed by John A. Hawkins, we may embed the discussion of definiteness within a general pragmatic model of speaker-hearer interaction (Hawkins 1978). In the interchange of information between speaker and hearer, one of the important factors is that the hearer should be able to know, with regard to noun phrases used by the speaker, whether he, the hearer, can already identify the referent of that noun phrase on the basis of previous discourse or current situation, or whether the referent is not (at least: not yet) identifiable. Many languages allow or require that the speaker leave cues in his sentences guiding the hearer in the solution of this problem, the referent identification problem. In English, the definite article *the* is an explicit signal from the speaker to the hearer that the speaker expects that both speaker and hearer can identify the referent of the noun phrase bearing the definite article; thus if an English-speaker says *Hasan bought the ox*, then he expects the hearer to know which ox is being talked about, and claims to know himself which ox is being talked about. This applies equally to the object noun phrase in Turkish sentence (34).

In sentence (35), although the referent of the direct object is not identifiable by both speaker and hearer (if it were, the speaker would not be asking 'which one?'), yet still the use of *hangisi?* 'which?' rather than *ne?* 'what?' implies that the set to which the referent belongs has already been delimited. More accurately: the use of *hangisi?* in

Turkish and of *which?* in English implies that there is a definite (in the strict sense) set to which the referent of the noun phrase belongs; this set is identifiable by both speaker and hearer, although they cannot both identify which individual member of the set is being spoken of. The reference is thus to an indefinite member of a definite set. Although the referent identification problem is not solved, it is partially solved, since the identifiability of the set to which the referent belongs significantly reduces the range of possible referents. Thus we might say that the so-called definite direct object marker in Turkish indicates that the referent identification problem is at least partially solved; in English, the definite article indicates that the problem is completely solved.

If we now turn to the contrast between Turkish sentences (33) and (36), both of which are glossed in English as 'Hasan bought an ox', although (36) contains the seemingly contradictory combination of indefinite article and definite direct object marker, then an even more specifically pragmatic solution is called for. In terms of (truth-functional) semantics, there is no distinction between (33) and (36), but once we ask about their roles in discourse, they turn out to be very different. Sentence (36), with the so-called definite direct object marker, would only be appropriate in a discourse if the referent of this noun phrase (in our example, the ox) is to recur subsequently in the discourse. If the ox is not to recur, for instance if the speaker is just listing Hasan's activities for the day, then only (33) would be appropriate. The function of the so-called definite direct object marker in (36) is thus to draw the hearer's attention to the importance of establishing the referent of the noun phrase; when introduced into the discourse, the referent of the noun phrase is not identifiable (whence the indefinite article), but the hearer is being told to bear in mind its reference for the purposes of the rest of the discourse. If the ox does not recur in the rest of the discourse, then he will feel that the speaker has cheated him. This use of the so-called definite direct object marker can be called: relevance of referent identification.

To summarize this section, the so-called definite direct object marker in Turkish serves to indicate that the referent identification problem is at least relevant and/or at least partially solved. The unifying factor is the referent identification problem, part of the pragmatic interaction between speaker and hearer. By refusing to be satisfied with the view that 'definiteness' can mean different things in different languages and to reject the possibility of finding language universals in this area, we are led not only to uncover the core that is common to 'definiteness' in English and Turkish, but also to provide a unified account of the definite direct object marker in Turkish. In this way, language-specific studies can sometimes directly benefit from work in language universals.

5. *Impersonal Passives.* So far, I have been arguing that work on language universals should take account of data from a wide range of languages in the establishment of language universals. I would like now to take this one step further, by claiming that statements about language

universals should be empirically testable (and thus potentially disconfirmable) statements about the possible occurrence of constructions across the languages of the world. This contrasts somewhat with the view prevalent within transformational-generative grammar, where universals are usually predicated of abstract underlying structures or of the abstract mechanisms required to convert these underlying structures into surface structures. A universal predicated of an underlying structure is no stronger than the validity of the underlying structure on which it is based, and given the immense range of disagreement over the nature of underlying syntactic structures, this basis seems very flimsy indeed. On the other hand, universals predicated directly on surface structures and relations among surface structures are much more readily testable. I shall illustrate this by contrasting two possible analyses of the so-called impersonal passive construction; this discussion is an abbreviated version of that presented in Comrie (1977).

In the English passive, as compared with corresponding active sentences, the subject of the active appears as an oblique object (with the preposition by) (as in (38)) or is omitted (as in (39)); while the object of the active appears as subject of the passive:

(37) *Norman hit Oscar.*

(38) *Oscar was hit by Norman.*

(39) *Oscar was hit.*

This kind of passive construction, with a referential noun phrase as surface structure subject, may be called the personal passive.

In a number of languages, however, in place of or in addition to the personal passive there exists an impersonal passive: in this construction, there is no referential surface structure subject (though there may be a dummy subject, e.g. like English *it*), and if there are any objects in the active these remain as objects in the passive. The following examples are from German, Latin, and Welsh, respectively:

(40) *es wurde gespielt.*
it was played
'there was playing'

(41) *acriter a militibus pugnatum est.*
fiercely by soldiers fought is
'there was fierce fighting by the soldiers' (literally: '(it) was fought fiercely by the soldiers')

(42) *lladdwyd y dyn gan ddraig.*
was-killed the man by dragon
'the man was killed by a dragon' (literally: '(it) was killed the man by a dragon')

A fairly traditional analysis of such constructions is to say that, in relating active and impersonal passive, the subject of the active is

either deleted or becomes an oblique object, while the subject position either remains empty or is filled by a dummy pronoun, such as German *es* 'it'. Proponents of Relational Grammar, in particular Paul M. Postal and David M. Perlmutter, have claimed that this analysis is incorrect, in particular that it violates a universal restriction against spontaneous demotion to oblique object: in other words, they claim that a noun phrase that is subject, direct object, or indirect object can only be demoted to oblique object if some other noun phrase is inserted into its grammatical relation, in which case the noun phrase formerly occupying that grammatical relation is automatically made an oblique object. With the personal passive, this gives no problems: in (38), *Oscar* has been promoted to subject, thereby displacing *Norman* to oblique object. Postal and Perlmutter's analysis of the impersonal passive, where apparently a subject has been spontaneously demoted to oblique object, is to claim that what actually happens is the following: a dummy noun phrase is inserted into subject position, and the insertion of this dummy noun phrase (just as of a referential noun phrase) demotes the original subject to oblique object.

This analysis has some plausibility for German, where in surface structure there is indeed a subject pronoun *es* 'it'. In Latin, although there is no subject pronoun, one could argue that the third person singular verb form (*est* '(it) is') contains a reflex of a third person subject, subsequently deleted as third person neuter subject pronouns usually are in Latin. The analysis runs into problems, however, with the Welsh example: in Welsh it is not possible to insert any subject in the impersonal passive construction, moreover the verbal form found in this construction is distinct from each of the six possible person-and-number combinations (e.g. third person singular *lladdodd* '(he/she/it) killed'). Thus in Welsh, there is no evidence in favor of positing any stage of derivation in which there is a dummy subject in the impersonal passive.

Let us now ask about the repercussions of the Welsh data for Postal and Perlmutter's putative universal banning spontaneous demotion. It seems to me that there are only two ways of reconciling the universal with data like these from Welsh. Either the universal is invalidated, since Welsh has an impersonal construction in which the subject is demoted to oblique object position without any noun phrase being inserted into subject position. Or, if we accept Postal and Perlmutter's analysis whereby a dummy subject is inserted and subsequently deleted (realized as null) without leaving any trace, then the universal is vacuous: there can be no possible counterexample to this universal, since it simply represents a decision on the part of the analyst to analyze all data in a certain way, and tells us nothing about possible sets of data. If the study of language universals is to have any empirical validity, then vacuous universals of this kind must be excluded.

6. *Conclusions.* In this lecture, I have argued that the study of language universals, contrary to the widespread view within transformational-generative grammar, should be based on the study of a wide range of languages. Going further, I argued that language universals should be predicated more directly of data and sets of data, rather than of abstract analyses, if the study of language universals is to

be an empirically testable undertaking.

In passing, I noted that this approach, though essential to the construction of a valid theory of language universals, does bring with it a methodological difficulty: ideally, the linguist working on language universals should both cover a wide range of languages and have detailed familiarity with each of the languages he uses. In practice, however, increase in breadth of coverage inevitably brings with it a decrease in depth, and vice versa. As my final point, I would like to suggest one possible way out of this methodological dilemma. Together with Norval Smith, I have devised a questionnaire, or rather framework, for providing a detailed description of the syntax, morphology, and phonology of an arbitrary language in such a way that the material contained in this description can be readily accessible to linguists working on language universals--this latter point is guaranteed by the uniform framework of presentation (Comrie & Smith 1977). The first language description following this framework, a description of Hixkaryana by Desmond C. Derbyshire, is to be published in 1979 to inaugurate the Lingua Descriptive Series. My hope is that the success of this series will help to make practicable a goal--the study of language universals based on investigation of a wide range of languages --which I believe to be so desirable as a matter of principle.

ABBREVIATIONS

ABS Absolutive; ACC Accusative; ACTNOM Action nominal; DEFDO Definite direct object; ERG Ergative; GEN Genitive; INSTR Instrumental; NEG Negative; NOM Nominative; O Object; S Subject; SG Singular; V Verb.

REFERENCES

Chomsky, Noam. 1970. Remarks on nominalization. In Readings in English Transformational Grammar, ed. by Roderick A. Jacobs & Peter S. Rosenbaum, 184-221. Waltham, Mass.: Ginn.

Comrie, Bernard. 1976. The syntax of action nominals: a cross-language study. Lingua 40. 177-201.

------. 1977. In defense of spontaneous demotion: the impersonal passive. In Syntax and Semantics 8: Grammatical Relations, ed. by Peter Cole & Jerrold M. Sadock, 47-58. New York: Academic Press.

------. 1978a. Definite direct objects and referent identification. Pragmatics Microfiche 3.1.

------. 1978b. Ergativity. In Syntactic typology: studies in the phenomenology of language, ed. by Winfred P. Lehmann, 329-394. Austin: University of Texas Press.

------ & Smith, Norval. 1977. Lingua Descriptive Series: Questionnaire. (= Lingua 42. 1-72.) Available from Ms. A. S. de Haas, North-Holland Publishing Company, 362-364 Herengracht, 1016 CH

Amsterdam, The Netherlands.
Derbyshire, Desmond C. 1977. Word order universals and the existence of OVS languages. Linguistic Inquiry 8. 590-598.
------ & Pullum, Geoffrey K. 1978. Object-initial languages. Paper presented at the LSA meeting, Summer 1978.
Dixon, R. M. W. 1972. The Dyirbal Language of North Queensland. New York: Cambridge University Press.
------. 1977. A Grammar of Yidiɲ. New York: Cambridge University Press.
Greenberg, Joseph H. 1966. Some universals of grammar with particular reference to the order of meaningful elements. In Universals of Language, ed. by Joseph H. Greenberg, 73-113. Second edition. Cambridge, Mass.: MIT Press.
Hawkins, John A. 1978. Definiteness and indefiniteness: a study in reference and grammaticality prediction. Atlantic Highlands, N.J.: Humanities Press.
Keenan, Edward L. 1976. Remarkable subjects in Malagasy. In Subject and Topic, ed. by Charles N. Li, 247-301. New York: Academic Press.

WHATEVER HAPPENED TO COMMUNICATIVE COMPETENCE?[1]

Susan Ervin-Tripp
University of California, Berkeley

0. *Psycholinguistic view.* The first formulations of psycholinguistics were derived from studies of rote learning, from information processing, and from Bloomfieldian linguistics, which was attentive to the language of reference. The social nexus of language acquisition and use was considered irrelevant to the study of language and unrelated to its structure. In the subsequent decade, the predominance of Chomskian perspectives in American linguistics guaranteed that autonomous syntax would be the center of attention in linguistics and allied fields.

1.0 *Communicative competence.* The term 'communicative competence' was first used by Dell Hymes to embrace all modes of competence in communication, to which speaking competence and linguistic code competence might be subsidiary. 'A child capable of any and all grammatical utterances, but not knowing which to use, not knowing even when to talk and when to stop, would be a cultural monstrosity' (Hymes 1967:16).

'The acquisition of competence for use, indeed, can be stated in the same terms as acquisition of competence for grammar. Within the developmental matrix in which knowledge of the sentences of a language is acquired, children also acquire knowledge of a set of ways in which sentences are used. From a finite experience of speech acts and their interdependence with socio-cultural features they develop a general theory of the speaking appropriate to their community, which they employ, like other forms of tacit cultural knowledge (competence) in conducting and interpreting social life.'

'In sum, the goal of a broad theory of competence can be said to show the ways in which the systematically possible, the feasible, and the appropriate are linked to produce and interpret actually occurring cultural behavior' (Hymes 1972:286).

Hymes was at pains to point out the distinction from narrower notions of competence and performance: 'The chief difficulty of present linguistic theory is that it would seem to require one to identify the study of the phenomena of concern to us here with its category of performance. . . . Indeed, language use is equated with performance: "the theory of language use--the theory of performance" (Chomsky 1965:9)' (Hymes 1972: 279). There was evidence that developing skill, errors, self-corrections, norms, and interpretations of violations applied not only to the growth of syntactic knowledge but sociolinguistic knowledge. It is an interesting coincidence that while a paradigm shift was being proposed from the perspective of a socially situated view of language on the part of anthropologists and sociologists, linguists and philosophers would

eventually push linguistics out again to include more semantic and pragmatic considerations.

1.1 *Field Manual.* This is the tenth anniversary of a significant event in the sociolinguistic revolution. In the sixties, the Social Science Research Council's Committee on Sociolinguistics sponsored summer workshops on Language, Society, and the Child in Berkeley. These were the outgrowth of an earlier informal seminar in which students and faculty participated including Hymes while he was still in Berkeley, John Gumperz, Dan Slobin, and I. Graduate students in the ethnography of communication joined us in producing a *Field Manual for Cross-Cultural Study of the Acquisition of Communicative Competence*, edited by Slobin in 1967. This was a first attempt to bring together the findings, types of issues, methodological proposals, and procedures which might arise if one took a 'communicative competence' perspective on the task of the child in developing language. Four dissertations grew directly out of that enterprise (Blount 1969, Kernan 1969, Mitchell-Kernan 1969, Stross 1969). In 1968 with the addition of Charles Ferguson we reexamined the issues and field results in workshops and seminars with 32 additional graduate students and a parade of visitors on the forefront of sociolinguistics (Ervin-Tripp 1969b). In this paper I propose to review the purposes of the *Field Manual*, briefly assess where we stand now on some major issues it raised, and look in some detail at problems in modeling the comprehension of situated speech--an issue which lies at the heart of the current convergence of the two historical streams I mentioned earlier.

The *Field Manual* set forth a specific series of purposes summarized in the introduction as follows:

a. 'The variety of systems in natural languages is so great that the study of language acquisition in varying languages can be regarded as a natural experiment which it would be extremely difficult to simulate artificially' (Ervin-Tripp 1967:ix). Were there universal generalizations to be made about the acquisition of syntax, phonology, and semantics? Since fundamental meanings are the same, what differences might appear because language provide differentially obvious maps for the communication of these meanings?

b. We wanted to know how changes in the social milieu might alter acquisition. We thought values about language and its uses might change how the child is spoken to and the types of speech promoted. There has been considerable controversy about the role of the social milieu in language development. 'It is not clear whether the order of acquisition of grammatical or sociolinguistic rules might be altered either by their place in the adult system or by specific values and practices of the child's caretakers. These practices include prohibiting or encouraging the child's presence when various speech events occur, and stimulating or responding to his speech' (Ervin-Tripp 1967:xi).

c. A more focused task in the same domain is the examination of the relationship between learning sociolinguistic rules and the learning of the basic code, or the impact of pragmatic context on code learning.

d. We wanted to examine the development of social competence and appropriateness in language. This was the most radical and innovative part of the program, for which there was least precedent in earlier work either with adults or first or second language learners. Children learn 'when to speak, when to be quiet, when to use ritual language, when to use baby talk, when to use polite forms, and when to shift language in a multilingual community.'

In addition, the choice of appropriate forms in sequence, and alternatives with proper social meaning, was to be included. 'To qualify as a native speaker in a speech community one must learn these rules as well as the rules of grammar. That is to say, of course, that one must learn to behave as though one knew the rules. Assessment of the acquisition of this competence should be as fundamental to the study of language development as the study of the formal grammatical code' (Ervin-Tripp 1967:x).

2.0 *Comparative grammar research.* Of these goals, the easiest to achieve would be the goal of discovering universals in grammatical development. From earlier work with a variety of (primarily) European languages we already knew a good deal about the development of syntax and phonology, the findings summarized in the *Field Manual*. We thus had tentative hypotheses based on the earlier experiments and diary studies, and a continually improving set of research methods.

Dan Slobin has been the principal advocate of the comparative approach, despite opposing claims that in-depth studies of English would accomplish more with less cost. He has pointed out that comparative studies of the same semantic structures can reveal what strategies children undertake when the available languages differ in the structural means they suggest for mapping (Slobin 1973).

2.1 *Word order research.* One important consequence of comparative studies is the disconfirmation of anglocentric theories of naturalness. It has been claimed that word order--a principal syntactic device in English--is so strongly dominant a natural strategy that it even precedes inflection in such languages as Russian and Finnish without support from the language spoken to the child. It has been claimed that English SVO (subject-verb-object) order, or at least SO, reflects the natural order of events in which agents exist before acts or objects. These assumptions appear to arise from the early use of telegraphic sentences in English.

If English order were easier, we would not expect to find Turkish children entirely ignoring word order as a potential syntactic device for understanding elementary semantic relations, as we have found they do. Nor would we expect to find Kaluli (New Guinea) children producing OSV sentences more often than SOV. Bambi Schieffelin however found that they were very common in family interaction. The argument before the verb is foregrounded in Kaluli, so that OSV is likely in claims, accusations, requests, and assertions about roles and rights, which are frequent in both adult and child speech in families (Schieffelin 1979).

If word order were easier than inflections, we would expect to find children who speak inflecting languages would be at a disadvantage in recognizing basic semantic relations in sentences. A large collaborative project directed by Slobin compared four languages. Turkish children do best, and Serbo-Croatian speaking children next best, in tasks requiring them to understand case relations at two.

The language spoken to Turkish children is, to be sure, not wholly random in word order; in the samples of adult speech, about half were in SOV order and a quarter SVO. But Turkish, as an agglutinative language, gives particularly clear and consistent mapping for case information. A child who consistently hears the same suffix for transitive objects can know that in the plural the same suffix will also be heard. And there is no gender to provide complexity, or in the case of gender of inanimates, an arbitrary disruption of case information.

In one study, children were presented with pairs of animals or dolls and asked to act out sentences such as 'squirrel scratches dog'. In Turkish these were presented in the six possible orders and with (O) or without (N) accusative suffixes. English children at two years of age essentially interpreted the various orders randomly. On Table 1 we see that Yugoslav children were able to interpret the inflectional information better than chance for NVO and NOV order but for no others, but Turkish children were essentially relying purely on suffixes and were the most successful. Where no suffixes were present (in the NVN, VNN, and NNV conditions) they did not differ from chance, displaying no order preferences, like the speakers of other languages at that age.

CORRECT OR SUBJECT-OBJECT INTERPRETATIONS
OF SENTENCES BY TWO YEAR OLDS

	TURKISH	SERBO-CROATIAN
NVN[a]	57%	62
NVO	77*	76*
OVN	74*	52
NNV	65	54
NOV	84*	71*
ONV	74*	61
VNN	50	36
VNO	76*	60
VON	77*	59

*Approximately statistically significant (Slobin in press).

[a]N indicates a suffixless nominal, O a noun with accusative suffix.

Table 1

2.2 *Question comprehension.* In another procedure with the same children, we asked them to answer questions. In this study, we sought to extend an earlier English finding (Ervin-Tripp 1970) which showed that there is a developmental order of ease in interpreting question words and

that error strategies were based on supplying sentence arguments missing from the questions. *What* and *who* questions were asked with two dolls, one on a vehicle, so that we could discover with one question confusions of case and animacy. *Who bumped him?* might be asked when a man in a car bumped a boy. The answer might be *car*, *boy*, *man*. In questions of the type *What pushed it? Who did she push?* and *What bumped him?* we found that case errors were common in English, Italian, and Yugoslav two-year-olds' replies, but not in the replies of the Turkish children. So again, we see that the agglutinative system, which supplies very clear, distinctive, and consistent marking for case relations, is learned early. We cannot simply assume the Turkish children were better at the task for other reasons, because they did make animacy errors.

These two studies confirm the basic principle that ease of learning depends on mapping simplicity. They suggest also that word order is more difficult than affixation, for case semantics. One hypothesis may be that agglutinative affixes are treated like lexical items, which are acquired easily by two if the meaning is not difficult.

2.3 *Universals.* The comparative studies, then, have disconfirmed some common assumptions about syntactic naturalness. They have shown that differences in linguistic structure result in differences in ease of acquiring language. But they have also revealed some cross-language generalizations beyond those known before such as similarity of early semantic relations expressed in sentences (Bowerman 1973, Brown 1973, Ervin-Tripp 1973, Slobin 1973). For example, the question comprehension study shows that questions are ordered in difficulty. In all languages studied, *when* questions are the hardest even when any temporal reply is accepted. *Where* questions, though structurally similar, are easier. The difference is one of cognitive development, of understanding of time concepts, rather than of language.

For the same reasons, in the four languages locational terms developed in a similar order. In all four languages *in*, *on*, *under*, and *beside*, which relate two objects, were learned first. *Between*, *back*, and *front* (for inherent back and front of houses and trucks) were next, and front and back of unoriented objects last (Johnston and Slobin 1977).

These comparative studies have been based on the assumption that the referential world to which children's earliest speech refers is common to the speakers of different languages. The difficulties experienced by children with respect to conceptual difficulties in apprehending time and complex spatial relationships will be similar. What differs then is the set of categories and mapping rules required to enter the semantic and syntactic expression system. In respect to fundamental cognitive development we can expect to find universals across language; the means of expressing those universals differ. The comparative studies reveal that languages are not equal in difficulty.

3.0 *Effects of milieu.* The studies of structural development were conservative in purpose. They were not especially affected by the social facets of the research program. There is a widespread belief that milieu

does not affect structure. Milieu supplies raw material for a robust innate gift for formal analysis. The innate language propensity is more sensitive to pattern than to frequency, to form than to reinforcement.

3.1 *Autogenous syntax.* The most dramatic support for this position came from a cross-modality study within the United States. Feldman and Goldin-Meadow located six deaf children whose hearing families were committed to an oralist education, to not consciously using natural gestures or sign language. Apparently doomed to noncommunication by their families, these children started with the spontaneous gestural system which is the nexus of all children's first language (Carter 1974) and each developed a consistent multigesture 'syntax'. They began to use pointing in combination to convey semantic relations between referents, and developed 'motor-iconic gestures that seemed to specify predicates of various kinds'. As the signs combined it became apparent that they were not random and not combined by sign type, but contained regular semantic sequences of the form patient-act, patient-recipient, or act-recipient. Patients were both first and most frequent. In the case of intransitive acts, the order was actor-act or referent-attribute (Goldin-Meadow 1975). These findings suggest both a robust innate propensity for formal regularity, and the low claim of English order on naturalness.

3.2 *Correlational research.* The strongest data on milieu effects on syntax have come from experimental and correlational studies within American families. Elissa Newport et al. (1975, 1977) collected syntactic samples from infants at two ages and compared changes in various measures to features of the mother's speech at the first testing. She found a series of finely tuned correlations which fit neatly with predictions about children's use of environmental information. Speech audibility was the strongest predictor of the rate of growth in English inflectional suffixes. *Yes/no* questions with modals, which front auxiliaries, brought faster rates of growth in the use of auxiliaries. In general, imperatives depressed language growth. Presumably the percent of imperatives is a reflection of very general differences in when and how language is used. Syntactic counts merely assess a consequence of higher-level decisions (Newport et al. 1977, Ervin-Tripp 1976b) to regulate, elicit, teach, and so on.

3.3 *Ethnographic description.* One consequence of cross-cultural studies has been the development of detailed description of the milieu of language use. As Ben Blount found in studying the Luo in Africa (Blount 1972) the formal data on child syntax may be strongly affected by the parental eliciting style and the constraints on possible talk settings for children. A vividly contrasted example of the 'discursive workshop' in Kaluli has shown the importance of elicited imitation as an instructional method and of cultural concern with interactional strategies.

In one Kaluli scene, two children eat food belonging to an absent sister and the mother uses the two-year-old as an agent to shame the two into putting the food back. Everyone knows whose food it is.

Mother to two-year-old in presence of other two children:

Whose is it? Say that.
Is it yours? Say that.
Who are you? Say that.

As instructions, these forms can be recognized as analogous to American mothers' directives such as *Whose are these?* or *What are those doing there?* involving out of place possessions. What the Kaluli add is the infant as an intermediary in these hints. In this way, parents and older children help toddlers accomplish identifiable social ends by telling them what to say to a second or third conversational partner. This is a technique of teaching contextualized discourse routines, showing children how to offer, request, question, assert, inform, challenge claims, threaten, tease, and shame (Schieffelin 1979).

Just as one-to-one mapping of referential semantics and form in an obvious, here-and-now situation has been shown to accelerate language learning, we can expect that learning of pragmatic contrasts might be easier this way. Contextualized modeling of specific formulae is a partial solution. Kaluli children are shown how to get what they want verbally. Sometimes they are also shown--as in the above example--how to get what their teachers want. Kaluli children do not have to extend formulae heard in the speech of others to themselves by guessing when the motives are the same. The Kaluli teachers do this extension for them. They tell the children exactly what to say. Berko Gleason and Weintraub (1976) have pointed out limited use in English of this method but merely for politeness formulae and restricted social routines.

From detailed ethnographic description, we hope to see how the children extend these samples to their own use and make the general patterns productive. In Kaluli the children's over-generalizations reveal clearly which aspects of the modeling have entered their competence. What we cannot yet know from such studies is whether such an instructional style leads to faster development of one aspect of competence over another.

4.0 *Pragmatics and syntax.* Does language use affect syntactic development? In the *Field Manual* it was pointed out that 'we know nothing about the relations between these aspects of learning and the learning of the basic code.' Studies by Halliday (1975), Ervin-Tripp (1977a), and unpublished work of Wells and Schieffelin have addressed these issues.

4.1 *Order contrasts.* The field study of Schieffelin showed that although Kaluli is typologically SOV, this order is used primarily in affectively neutral narrative occasions and is less frequent in families than OSV which focuses on the agent. So important is this contrast that a child's demand in SOV (apparently generalized from a preceding narrative passage) was ignored by the mother. The same demand in OSV usually was not.

An order contrast, then, is a pragmatic difference, as it probably is in many languages. It is learned very early. Eventually the children add an agentive suffix, which is obligatory in the OSV order, but which the children at first overgeneralize to SOV (but not to intransitives). Thus such syntactically important developments as agentive marking occur

in the context of a prior pragmatic contrast.

4.2 *English production rules.* Is syntax integrated with pragmatic development in English? In a study of syntactic development in English two-year-olds I have sought production rules for each child (Ervin-Tripp 1977a). The purpose was to account for the telegraphic character of these utterances, the omission of semantically important arguments, diversity of order. Those who have worked for years with child texts recognize a high degree of predictability in child utterances, which production rules should capture. This predictability is obviously related to context and purpose.

My tactic was to start with the larger context and to seek regularities for achieving conversational strategies. An example is the analysis of local reply rules. Sally at 1;43 (one year, 43 weeks) had reply patterns as follows:

Routines: Bye-bye Sally: *Byebye Sally*
How are you? / Where are you? : *Fine*

To questions with rising pitch:
- *Yeah*
- REPEAT (stress changed segments)
- Nonresponse
- (for any of above carry out feasible act mentioned by other)

Hm: SELF REPEAT (added at 1;45)

This description displays the fact that children's compliance, or not, with explicit directives may be accompanied by replies which are responsive to the rising pitch in the question format of directives such as *Can you open the door?* There isn't necessarily a congruence between the two response systems. (See below, 5.4.3.)

A more elaborate description of a reply and initiation system of an older child (2.3 to 2.5 or nearly two and a half) shows considerable enlargement of abilities. This child differentiated between directives and information questions in replies, and could herself initiate a question. Analysis being of natural family interaction, some events did not appear in the corpus.

This analysis by categories has a strong debt first to Michael Halliday and then to Gordon Wells, who worked out functional categories (Wells 1973). The method was to make an analysis by function or act, and then make a matrix which matched form against act. The description in Table 2 sought the most concrete formulation which minimized postulation of either function or syntactic categories. Examples of the classes labeled proper name, act, object of action, and common object are these which were grouped in the following chart as ACT (+ OBJ): *put it, tear this, do cape, eat my breakfast, see, come-here,* and *go-peepee.* While the grouping cuts across transitive/intransitive, the pragmatic categories, such as 'announce achievement' did too.

PRODUCTION RULES OF ELLY (2.3-2.5)

Replies:
(to directive) Assent: { *O.K.* / *Kay* } Refuse: { *No* / *unh-unh* }
(to rising pitch) Agree: *yeah* Deny: *No*
(to statements) Affirm: { *yeah* / REPEAT }
(to insult) Epithet: REPEAT
(to question) Where question: { *There he/it is.* / *There's* { NAME / OBJECT } }

Initiatives:
Call attention: (*Hey*) + NAME
Command Observe: { *Look* / *Lookit* OBJ }
Offer: { *You like it?* / *Want this?* / *Here* }
Claim: (That's) { *mine* / *my* OBJECT / NAME'S }
Prohibit or Mock Warning: { *Stop it* / *Don't* }
Permission request: *Can I* + ACT?
Command: (*you*) + ACT (+ OBJ)
Complete act: *There*
Announce achievement: *I* + ACT (+OBJ)
Solicit info: *What's that?*
Identify: *That* OBJECT

Table 2

In this proposal, functional development is treated as primary. The studies of gestural development, such as Carter's (1974), suggest that there are some intent primes, just as there is evidence from such studies as Goldin-Meadow's (1975) of such cognitive features as patient--act of transfer--recipient. Starting with function categories resulted in a massive simplification of the data in early speech especially in natural interaction where interpersonal acts are diverse and speech is often formulaic. Failure to make such classifications is likely to result in confusing superficially parallel instances in which the production system is quite different, as in *I want dolly* and *I build tower*. It became apparent in this work that the syntactic rules one can derive from textual data depend strongly on the type of interaction sampled. We found extreme differences between peer speech and the speech with adults who use special speech-eliciting acts with children.

When contextual, functional, and discourse information is included in production rules, the level of abstractness of syntax may prove to be fairly limited at first. When there is evidence of transfer of patterns across function categories, the degree of generality of syntactic categories may increase until syntax must be treated as a separate system. It still could remain the case that for some functions, such as greetings, production rules could remain relatively low level and formulaic. In the case of directives, forms such as 'Can you . . .', which begin as formulas and are commonly so used, may be reanalyzed later to permit production of such paraphrases as 'Would you be able to . . .' Thus the level of abstractness may change even in adults.

Just as children appear to differ in whether their pragmatic

competence is elaborated more than their referential system in speech (Dore 1973) it seems likely that social emphasis on pragmatic differentiations such as the attention in Kaluli to social strategies might affect the developmental sequence in the emergence of syntax.

5.0 *Sociolinguistic rules.* The discussion of sociolinguistic aspects of competence in the *Field Manual* included style-, register-, and code-switching, politeness, elaborated and restricted code use, the contextual factors which condition formal alternation in speech, and the sequential structure of speech events. Collections of studies of these topics have appeared in Ervin-Tripp and Mitchell-Kernan's *Child Discourse*, Ochs and Schieffelin's *Developmental Pragmatics*, and in *Language in Society* Vol. 7, No. 3, 1978. Only a few of these studies were deliberately comparative, specifically the studies of directives by Hollos and Beeman (1978) and of arguments by Lein and Brenneis (1978).

5.1 *Mapping complexity.* When we address someone, or make a request, we do at least two things at once. We accomplish a communicative act such as calling out, or requesting, and we also convey our social categorizing of them and of our relationship. If either is already known, the choice is predictable, there is no new information, and the listener can leave in the background this aspect of the interaction.

When an act is expected or obvious, the formal alternatives can focus on the social relationship. In a study of roommates, a roommate said to one blocking his way 'You make a good door, Sal'. Since his purpose was sufficiently obvious, his focus could be on a solidary joking relationship. In such cases, elliptical reference, such as mention of the door, was adequate. If both the social relationship and the act are unclear, several alternative interpretations can compete.

INTERPRETATION AND CONTEXT:
Child-adult dyads

	Adult to Small Child		Child to Adult	
Resource Access:	A	C	A	C
D'you wanna play on the swings?	Offer Directive Info. Req.	Directive Info. Req.	Info. Req.	Offer
Can I help you?	Offer Declare Intent	Offer Declare Intent	Permission Req.	Offer Permission Req.
Take some	Offer Directive	Directive	Directive	Offer Directive
Can I have some?	Mock Permission Req. Information Req.	Directive	Directive	Permission Req.

Table 3

In Table 3, the same words can accomplish different acts or have different purposes, and the same acts can be expressed in different ways. We have this complex system because it allows us to do several things at once very efficiently. In the case of calling out to people, it is clear what we are doing. Consider the following example: One of my children answered the telephone briefly and then shouted *Hey ugly*! The act of summoning was obvious. The insult was obvious. What was left unspecified was who was addressed. I knew that it wasn't meant for me. The insulted person had to identify himself and to volunteer to accept the insult.

Research on social meaning in the sixties focused on the social complexity of address rules (Brown and Gilman 1960, Ervin-Tripp 1969a), and it appears likely that some universals can be found in address systems. These studies typically hold constant and do not attend to the interactional or discourse functions of the naming. They assume function is not a problem. When we examine communicative acts such as directives we add two new types of complexity. One is that we are dealing with sentences, not just noun phrases, so that there is much more structural complexity. The other is that the purpose can be hidden. The problem of recognizing purpose has attracted attention from philosophers and linguists since they cannot compute conveyed intent easily without additional information beyond syntax and semantics. What has not been seen in this research is its intersection with studies of social alternations. Table 3 illustrates that social meaning and understood intent must be interconnected systematically.

5.2 *Contextual prediction.* Recently at a family dinner, I heard a niece, looking at the home-baked cake before her at the center of the table, say *Is the cake there?* A typical analysis of my understanding of this utterance would require that I entertain and reject the hypothesis that she wanted to know if the cake was there as a request for information about location. Having rejected that interpretation I could turn to other possibilities for her intent.

We are all aware, at times, of such searches. But there must be a difference between the occasions on which we are aware and those in which there is immediate rapid interpretation. Also, there must be a reason why we must, to ensure a literal interpretation of such a question, make sure that the more routine intent is not heard. Such occasions are numerous, as when we just want to know if someone is there and not have them summoned to the phone (Ervin-Tripp 1976a). Such an interpretation will not account for how children acquire communicative competence in hearing imbedded directives such as *Do you wanna play on the swings?* before their syntax or computational inference skills are fully developed. Such a view omits the state of the hearer at the time before the utterance occurs. I wish to argue that most of the time utterances do less work than we have assumed. I will also argue that it is precisely this fact that allows us to develop the potential for understanding the remaining cases.

We have been examining nearly a hundred videotaped scenes with families, seeking attempts to control conversational partners. The families were videotaped in their homes, and they included children ranging in age

from one to eight years. In some scenes there were siblings, in some visiting friends. In these videotapes, we noticed a strong sequential structuring of activity. The activity sequences might include initiation, negotiation, stage-setting, internal structure according to the activity, repairs, episode units, and terminations. Sometimes we saw the children doing a lot of work to set the scene for an activity that doesn't take place. They allocate roles, place people and objects in space, allocate or claim territory. Sometimes there are structured or routinized activities such as games. If we observed discussions around topics, the topical sequences would have internal structure.

At the same time, we can see orderly sequences in surface structure. In a text we may find boundary markers, turntaking, style shifts, and such local sequencing as repetition, replies, and tying forms. For example, questions solicit answers, people pronominalize or form ellipses based on previous utterances, add on with conjunctions. These features can, to some extent, be examined without regard to the properties of activity.

Activities may have very strong structure. In a game of lotto, when the mother picks up a card it must be claimed and placed. Her speech then is quite redundant with prior knowledge of the participant, if she says *Where does this go?* and shows the card. If she says *Who has the bear? How is the bear?* or *Here's the bear* she identifies the bear instead of showing it, so the net effect is the same. From the standpoint of activity consequences the interpretation of all these surface forms is identical.

At the same time, these sentences differ in their surface sequential properties. If she says *Where does this go?* there might be a greater chance of eliciting *Here*. If she says *Who has the bear?* she might more often receive *Me*, but we can expect considerable overlap in replies because replies may be more attentive to the activity system than to the surface structure.

Realization rules display purposes through the choice of words or actions. In this example, the realization rules would have to show that there are several ways to make offers in a lotto game. We don't know why the mother chooses different formulations at different times.

Realizations, as we saw in Table 3, are affected by social information. Social information includes both givens with which participants enter the scene, like age, sex, power, and setting, and negotiable social components like affect, task, and key or mood. Listeners who have little prior information may be able to guess power, feelings, or task from the choices, so that social reality and speech influence each other.

Because of the effect of information available at a given point in discourse, which includes activity or topic, the sequence of surface forms, and social givens, it is possible to construct predictions about what will happen next. We recently tried this with a videotape from our corpus, giving viewers a transcript and letting them see both transcript and scene up to a critical point, and then moving to another test point. Overall they were successful half the time in predicting the next act.

They were surprised at their own success, which sometimes included verbatim accuracy.

5.3 *Act and form.* The relation between activity and realization is affected by the intention of actors, which can of course alter tasks. In looking for the relation between intention and verbal forms, we were faced with the formidable task of identifying speaker purposes. Formidable, yes, but we do it all day in natural interaction, so this skill must be well-developed in sociable humans. To delve deeply, as in the study of Labov and Fanshel (1977), may be more difficult. So we decided to begin at face value and look for formal realizations of intent.

Our categories for intent were directed at attempts to control only the acts of others, not their beliefs, though we will see that directing attention may be a primary means of controlling acts.

The list we made could be criticized, but it is a beginning. We have been able to get acceptable agreement among judges in these assessments. Acts are split between those which seek to control the hearer's actions, and those which by defining the speaker's actions indirectly constrain the hearer. In the first set are directives, prohibitions, warnings, and ownership claims. In the second are offers or promises, asking permission, and asserting intentions. In the texts we have examined so far most control acts are directives, fewer being offers and ownership claims.

LIST OF ACTS TO BE CODED

DIRECTION to perform some action, render services, give goods.
E.g., Is there any cake left?
PROHIBITIONS to prevent the addressee from acting, or to keep out of territory.
Stay in your own room.
THREAT of action on the other unless there is compliance.
I'll tell Mom.
WARNING to call attention to a danger or problem which will affect the addressee and which requires prompt preventive measures.
Mom's gonna kill you.
OWNERSHIP assertion of status or territory, including prohibiting use by others.
That's my chair.
OFFER or PROMISE of goods and services which will have some benefit to the recipient.
I want to do that for you.
PERMISSION/REQUEST for the speaker's action when the beneficiary is the speaker.
May I please go out in the rain?
INTENTION by the speaker to commit an act affecting the hearer or hearer's interests, whether or not the hearer replies.
Imply directives or prohibitions.
I'm gonna sit on your lap.

Table 4

The primary formal categorization we have made is a division between explicit and implicit forms as seen in Table 5.

SPEECH ACTS CODED

Form	Example	Possible Intent
Gesture or vocalization only.	(Show food.)	Offer.
Ellipsis, explicit in context.	Me too!	Directive.
Imperative.	Quit that.	Prohibition.
Explicit question.	Can you stop!	Prohibition.
Permission questions.	Can I have some?	Directive.
Permission statement.	You can go outside now.	Directive.
Explicit statement.	That's mine.	Claim.
Aversive state.	I'm hungry.	Directive.
Implicit condition.	It's finished.	Directive.
Implicit question.	Is there more?	Directive.
Normative statements.	Only mommies say honey.	Prohibition.

Table 5

You'll notice that any of these forms might be used for a particular intent. I could point at a glass of wine, say *me too*, *give me some wine*, *could you give me some wine*, *can I have some wine*, *you can give me more wine now*, *I want some wine*, *I'm thirsty*, *I'm out of wine*, *is there more wine?*, or *Californians always have two glasses of wine*. These alternatives, of course, carry different consequences for discourse (Ervin-Tripp 1976a).

The first texts we have coded have shown some differences in use in families; the analysis was drawn from 2500 utterances, of which 174 were child control acts.

A. We found many more requests and directives than other types of controlling acts in our sample. Over half were expressed as imperatives. Ellipsis was also very frequent.

B. There is a clear change in repertoire with the age of the speaker. At one, we found cries, gestures and one-word directives like *no*. At two, need statements and imperatives were added. At three, children add permission questions and explicit embedded questions like *can you give me that?* which require some formulaic frames. By four we saw the first explicit statements, or hints, but they were rare and used with adults.

C. There was confirmation of the nursery school evidence (Ervin-Tripp 1977b) that by two and a half children already differentiate forms for different addressees. To other children they produced imperatives and ellipsis. To mothers they produced imperatives, ellipsis and need statements. To adult outsiders, they gave permission requests and explicit questions but not imperatives or need statements.

5.4 *Testing understanding.* The directives we saw were typically redundant with gestures. So we began to design experiments in which it would be possible to alter contextual features systematically. In this series of pilot studies, we wanted to change the situation and what people say separately. So we made cartoon stories which allowed us to add or remove pictures. We have tape recordings on cards which allow the exact features of the stimulus to be controlled.

EXPERIMENTAL SENTENCES VARYING IN EXPLICITNESS

1a. Oh, my green pen.	1b. I don't like my green pen
2. Is my green pen there?	
3. My green pen's there.	
4. I can't find my green pen.	
5. Can you find my green pen?	

Table 6

In the cartoon stories we worked only with middle-level explicitness. For example, if two children are fighting, the form might be *Are you fighting?* or *You're fighting.* In the case of helping stories, the form might be *Is the checker there?* or *Suzy, the checker's there.* We tried variations in intonation, and compared whether a child or mother spoke to the two children in the pictures, whether there was a visible speaker who could see what was going on, and whether no one spoke at all. And we changed two other contextual factors--whether they were doing something forbidden, or simply were present in a case where there was a problem requiring help.

How can you tell what a three-year-old understands when such a story is told? We asked what the children in the story did, what they said, and why. If these produced no clear solution, we proceeded to ask what they would do in a like situation, and what the mother or third child wanted.

5.4.1 *Practical reasoning.* At three, the children are quite clear about what acts are forbidden, and know that they should stop fighting, throwing food, trampling on the flowers, and getting mud on their party clothes. But they stop primarily on the basis of who is present, either in sight or sound. If the mother can be seen or heard, they stop in 87% of the cases. What is said is irrelevant, since they stop just as often for a silent mother. The children know the rules of action, and when the enforcing agent is around they obey. But they are not reacting to orders though they give the impression of understanding the directives.

5.4.2 *Effects of verbal form.* In the problem situations, on the other hand, the three-year-olds rarely supplied help without being guided by speech, so speech was informative. Since the particular speech form used in the stories didn't matter, it appeared that just referring to the problem helps at each age. In order to check more closely on the effects of form on helping, we developed some real-life situations with lost pens

and falling objects. At various times during the picture-story procedure we reminded the children, using one of the forms on the continuum in Table 6. For the most explicit form *can you . . .* even the three-year-olds helped in 90% of the cases. For statements, they helped about half the time, but for the questions and naming they were rarely successful.

By five, form makes no difference, in real life helping, even in some second-language learners of English. Even incongruous mention of a desired object, as in 1b on Table 6, is effective, provided the speaker is hunting for something. So for the smallest children, the more explicit the form, the greater the compliance. By five practical reasoning is sufficient so only a mention is adequate.

5.4.3 *Reasons for action.* In these situations, we can get four kinds of information about what the children know. They said what was done, what was said, why and what the speaker wanted. What is the relation between these? The orthodox interpretation of speech acts suggests that the listener infers what the speaker wants from the form of the act, and then checks the situation to see if that could be true. The imputed order of processing is analysis of speech, interpretation of intent, check of context, and then behavior. Where a meaning must be discarded, the interpretive process must be repeated.

Three-year-olds, of course, do not easily discuss intent. But here is a nicely expansive reply:

(When the mother and child come to the door with groceries, they ask Jack and Kate, 'Is the door open?')
--What do Jack and Kate say?
--I don't know
--Well, what would you say if your mom was carrying groceries and she said Jenny, is the door open?
--Yes. (It is closed in the picture)
--You'd say yes?
--But I'm going to open it.
--Oh, you'd open it!
--Yeah, but the wind gonna blow it closed again.
--Well, why would you open it?
--Cause I want them to get the groceries in the house so I can eat.

This reply doesn't mention what the mother says or wants. The grocery scene involves a normal activity sequence in which groceries are brought from the car into the house through an open door. The mother's mention of the door has drawn attention to a barrier in the normal activity sequence. The child, wanting the sequence to be carried out in full therefore opens the door. But the motives of the mother's remark or what the mother wants are not part of the inferential sequence.

By five the children can reason a good deal about the motives of the children in the stories. Over and over they told us that the reasons for their actions were practical. In the scene with the checker, the mother said *I can't find the checker.* The child replied *Here it is* and gave it back. Why? *Because she needed it.* When the children do not

supply help which is obviously needed, the reason is because they are mean.

Does the answering of the mother's question play no part in the process leading to action? Some analysis of the understanding of speech acts see this as part of the reasoning process. Of course the form of the reply has a surface relation to the question. In the real life directives, a third of the *Can you . . .* forms elicited replies, but *Is X there?* which had less likelihood of leading to action, was answered by two thirds of the children. There was no relation between whether the children replied, and whether they acted.

In the picture stories, too, there was no relation between replying and acting. The child who was asked *Is the checker there?* might say *yes* and give it back, say *no* and give it back, or say nothing and give it back. Or might withhold it and say *yes*. Thus it appears that action and reply are not related and may be incongruent.

Our evidence so far suggests that there are two parallel systems. One is a semantic system which checks truth value for questions and comments and leads to verbal replies. The other is an action system in which words affect attention and practical reasoning leads to actions.

5.4.4 *Speakers motives.* What the child knows can be judged from a variety of standpoints--from the effects of speech on what the child does, and from the effects on what the child thinks the speaker wants. Ultimately, from the standpoint of the theories of speech acts, the critical issue is what a listener thinks a speaker intended--the conveyed intent. So the most important test of children's knowledge should be how they answer the question: *What did the mother want, what did she mean when she said . . . ?*

We did not ask this question in the real-life solicitations for help from the clumsy experimenters who kept dropping and losing things, since such questions would have, we thought, been too intrusive in the experiment. So in what follows we shall be considering attributions of motives to speakers in the narratives. In half the conditions, the speaker could see what was going on, and the children said so. So from the standpoint of communicative economy, speakers should not be seeking for information in these cases. Yet a quarter to half the time (depending on the action depicted), the children said nevertheless that the speaker *just wanted to know, yes or no*, when she said *I can't find the checker, is the checker there?* or *are you fighting?*

In attribution of motive, the children were affected by the status of the speaker: mothers were more likely to want naughty acts to stop than were child speakers. Some of the children, when asked about motive, scrutinized the picture to check whether the speaker could see. When mothers could not see, they were more often assumed to be ignorant, and more often the seven-year-olds would plan to lie. Some very strategically sophisticated children, then, considered the point of view and state of knowledge of the mother right away; when asked what the children would say, they said they would lie. Then they could avoid trouble. The

children at seven were from the standpoint of assessing knowledge more likely to think of and be able to guess when the mother was ignorant. The three-year-olds, you will recall, assumed that even if the mother was out of sight she could see what was going on. This difference in ability to imagine the physical perspective of another is a well-known feature of cognitive maturation.

We could argue also that ability to imagine the motivation of others changes with age. In the case of the hints, the seven-year-olds vary. About half were immediately ready to attribute directive purpose to hints, even when the speaker was not able to see. Some said *The door's closed. Would you like me to open it for you?* The few teenagers we have questioned routinely assume directive motives in these cases, although in the case of the mothers, hints are particularly ill-formed ways of getting naughty acts to stop, and the children were able to offer much more likely explicit forms.

Conventional models for interpretation assume that the listener hears, computes a possible intent, assesses context, and then decides what to say and do. This model is clearly wrong for most seven-year-olds. When they were asked what the children reply, they reply equally for cases where they later say the speaker wants to know, and when she wants action. They reply equally to questions and statements (unlike the three-year-olds who reply little, and scarcely ever to statements). When asked what the children do, their actions appear unrelated either to their replies or their later attributions of motive. Over a third of the children who propose help later say the speaker just wanted to know. We had a strong impression from many of the children that the calculation of intent took place after decisions about replies and actions, merely because we asked about it. These children delayed, studied the pictures, sometimes re-assessed what they said earlier on the basis of probes--at least they recognized an inconsistency. Only the children who proposed lying were clearly assessing both perspective and possible intent at the time of initial replies. And in the condition in which the mother's voice sounded angry compliance was more common--suggesting a recognition of mother's mood as an early process, when it is a marked feature of the voice.

5.4.5 *Comprehension processes.* The experimental situations, as well as our work with videotaped natural scenes, suggest that what we have called activity sequences have a strong impact on what children expect or plan to have happen next. In the real-life hints from the experimenters, and in the stories with hints for help, the older children were eager to give help, to be cooperative. In these situations verbal form made no difference, since drawing attention to the problem was all that was needed.

In addition, speech to children can stop a competing activity, draw them into a new focused engagement so that providing active help is more likely. The three-year-olds who replied (usually to questions) in the narratives also were more likely to offer help. For the three-year-olds and half the seven-year-olds, the hints were not heard as expressing wants. They were interpreted literally, but they shifted the attention of the children to interaction with a new partner and to awareness of a problem.

In the stories about naughty actions, the children in the narrative were already engaged in activities which many of the older children admitted were lots of fun. The three-year-olds were especially delighted with playing in the mud and seldom proposed stopping! Since the rules were known by all, it is routinely assumed that mothers, at least, will want the children to stop if they know what is going on. We found the older children had more strategies for, and much more often proposed non-compliance. But most important, despite this strong knowledge of mother's values the hints were often not heard as directives, and still given literal interpretations by almost half the seven-year-olds.

We have argued against a model which assumes that listeners at any age regularly hear an utterance and then begin a very rapid computational process including assessment of motive. Our evidence suggests that:

a. Participants often do not assess motive at all.

b. A considerable regularity or predictability in activities derives from knowledge of usual or normal sequences. When verbal interchanges occur, they are redundant with what is already known, in many cases.

c. The new information provided by speech may be as little as mood, attraction of a listener to a new focused engagement, or attention to an unnoticed feature of the environment.

d. The system of local cohesion provided by questions and answers, comments and replies can be relatively independent of, and occasionally even incongruent with the simultaneous negotiation of activities.

Since it is obvious that competent adults are capable at times of calculating intent, and basing actions and replies on a multilevel analysis that depends heavily on verbal processing (e.g. Labov and Fanshel 1977) we have to account for how this capacity develops. The argument here is that the contextually rich situations in childhood and later are instructional instances. In the family videotapes we have repeatedly seen adults and older children redundantly producing directives and hints when children have no choice at all, when the next action is already under way, entirely predictable, or displayed with a gesture. We expect there may have to be four conditions for children to learn to hear novel hints as directives:

a. The child has to have a sufficiently adept linguistic capacity to interpret speech very quickly and make additional inferences while replying.

b. The child has to be socially mature enough to assess the intentions of others while interacting, that is, to take the view of the other.

c. Hinting must be a part of the directive repertoire of those with whom the child interacts, and must be heard in unambiguous circumstances. (This condition parallels prerequisites for the

acquisition of grammar as in Ervin-Tripp 1973 and Slobin 1973).

d. Possibly, the child must be able to impute to others face-protecting motives such as avoiding imposing, for certain types of social relationships. These appear to motivate the choice of hints, and to define when they are an appropriate or inappropriate strategy.

6.0 *Conclusions.* The diversity of work in the decade since the *Field Manual* does not lend itself to summarization. For most linguists, the study of child competence seems marginal at best. It is usually assumed that adult knowledge is static, fully evolved, not affected by new learning, not sustained by unambiguous contexts, and centered in a richly abstract and inferential understanding of language. The image one gets is of the infinitely wise scholar in his study, out of contact with a real world, able to understand any sentence out of context. The study of child acquisition compels us to recognize the other side of our adult abilities: that they are often used in repetitive, highly predictable circumstances, that they are far more efficient than complex calculations should permit, that they constantly are affected by new learning, and that their most abstract and context-free occasions are nurtured by the learning conditions of language used when we already know the structure of the activity, social relations, or ideas to which language refers. Thus the sources of language take us out of language.

NOTE

[1]The experimental work on pragmatics reported here was sponsored by NIMH Grant 26063. The data were collected by Georgette Stratos, Catherine O'Connor, Hannah Kaltman, and Amy Strage, whose work and discussion are gratefully acknowledged. The videotaped project was a collaboration with Jenny Cook-Gumperz and David Gordon.

REFERENCES

Berko Gleason, J. and S. Weintraub. 1976. The acquisition of routines in child language: 'Trick or treat'. Language in Society 5. 129-37.

Blount, B. G. 1969. Acquisition of language by Luo children. Unpublished doctoral dissertation, University of California at Berkeley.

------. 1972. Aspects of Luo socialization. Language in Society 1. 235-48.

Bowerman, Melissa. 1973. Early syntactic development. Cambridge: Cambridge University Press.

Brown, R. 1973. A first language. Cambridge, Mass.: Harvard University Press.

Brown, R. and A. Gilman. 1960. The pronouns of power and solidarity. Style in language, ed. by T. Sebeok, 253-76. Cambridge, Mass.: MIT Press.

Carter, A. 1974. Communication in the sensori-motor period. Unpublished doctoral dissertation, University of California at Berkeley.

Chomsky, N. 1965. Aspects of the theory of syntax. Cambridge, Mass.: MIT Press.

Dore, J. 1973. The development of speech acts. Unpublished doctoral dissertation, City University of New York.

Ervin-Tripp, S. M. 1967. Introduction. Field Manual for cross-cultural study of the acquisition of communicative competence. Berkeley: Language-Behavior Research Laboratory.

------. 1969a. Sociolinguistics. Advances in experimental social psychology 4, ed. by L. Berkowitz, 91-165. New York: Academic Press.

------. 1969b. Summer workshops in sociolinguistics: Research on children's acquisition of communicative competence. Social Science Research Council Item, Vol. 23, 22-26.

------. 1970. Discourse agreement: How children answer questions. Cognition and the development of language, ed. by J. R. Hayes, 79-107. New York: John Wiley and Sons, Inc.

------. 1973. Language acquisition and communicative choice. Stanford: Stanford University Press.

------. 1976a. Is Sybil there? The structure of some American-English directives. Language in Society 5. 25-66.

------. 1976b. Some features of early child adult dialogues. Language in Society 7. 357-73.

------. 1977a. From conversation to syntax. Papers and reports on child language development, No. 13, 1-21. Department of Linguistics, Stanford University.

------. 1977b. Wait for me, roller-skate. Child discourse, ed. by C. Mitchell-Kernan and S. Ervin-Tripp. New York: Academic Press.

Goldin-Meadow, S. J. 1975. The representation of semantic relations in a manual language created by deaf children of hearing parents: A language you can't dismiss out of hand. The acquisition of linguistic structure, unpublished doctoral dissertation, University of Pennsylvania.

Halliday, M. A. K. 1975. Learning how to mean: Explorations in the development of language. New York: Elsevier North-Holland.

Hollos, M. and W. Beeman. 1978. The development of directives among Norwegian and Hungarian children: An example of communicative style in culture. Language in Society 7. 345-56.

Hymes, Dell. 1967. Models of the interaction of language and social setting. Journal of Social Issues 23, No. 2. 8-28.

------. 1972. On communicative competence. Sociolinguistics, ed. by J. B. Pride and J. Holmes, 269-93. Baltimore: Penguin Books.

Johnston, J. R. and D. I. Slobin. 1977. The development of locative expressions in English, Italian, Serbo-Croation, and Turkish. Papers and reports on child language development, No. 13, 134-47. Department of Linguistics, Stanford University.

Kernan, K. T. 1969. The acquisition of language by Samoan children. Unpublished doctoral dissertation, University of California at Berkeley.

Labov, W. and D. Fanshel. 1977. Therapeutic discourse. New York: Academic Press.

Lein, L. and D. Brenneis. 1978. Children's disputes in three speech communities. Language in Society 7. 299-324.
Mitchell-Kernan, C. 1971. Language behavior in a black urban community. Berkeley: Monograph of the Language Behavior Research Laboratory, No. 2.
Newport, E., L. R. Gleitman, and H. Gleitman. 1975. A study of mother speech and language acquisition. Papers and reports on child language development, No. 10. Department of Linguistics, Stanford University.
------. 1977. Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. Talking to children: Language input and acquisition, ed. by C. Snow and C. A. Ferguson, 109-50. Cambridge: Cambridge University Press.
Schieffelin, B. 1979. How Kaluli children learn what to say, what to do, and how to feel: An ethnographic approach to the study of communicative competence. Unpublished doctoral dissertation, Columbia University, New York.
Slobin, D. E. 1967. A field manual for cross-cultural study of the acquisition of communicative competence. Berkeley: Language-Behavior Research Laboratory.
------. 1973. Cognitive prerequisites for the development of grammar. Studies of child language development, ed. by C. A. Ferguson and D. I. Slobin. New York: Holt Rinehart, and Winston.
------. In press. Evidence for syntactic change in child speech: Word Order in Serbo-croation and Turkish. International Review of Slavic Linguistics.
Stross, B. 1969. Language acquisition by Tenejapa Tzeltal children. Unpublished doctoral dissertation, University of California at Berkeley.
Wells, C. G. 1973. Coding manual for the description of child speech. University of Bristol School of Education, Bristol, England.

LEXICAL REPRESENTATION AND THE ACQUISITION OF PHONOLOGY*

N. V. Smith
University College, London

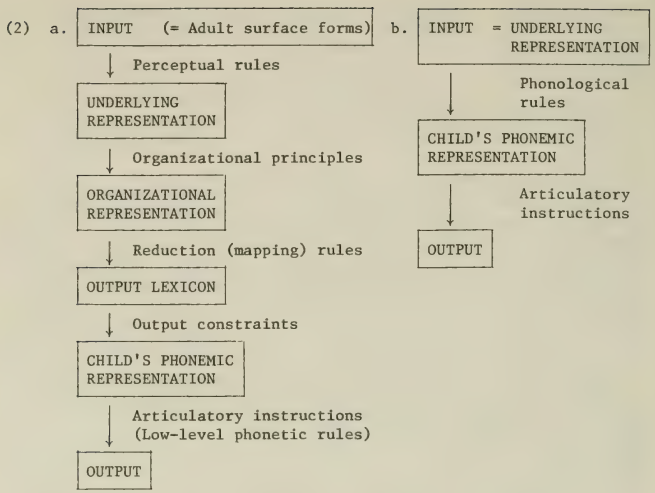
It is a truism that the pronunciation of young children differs from that of their adult models.[1] The reasons for this divergence are unfortunately not so obvious, and there is corresponding disagreement in the literature on what constitutes an adequate model for child phonology in general and the child's lexical representation in particular. Here I want to survey some of the evidence pertaining to a number of interconnected areas of dispute. I shall concentrate on the interplay of perceptual, productional, and organizational factors underlying such divergencies, in order to determine the nature of the child's lexical representation: how it differs, if at all, from the adult form he hears; and how it is related to what he produces: in other words, what constraints one can put on the mapping from lexical representation to output. This entails deciding whether it is necessary to postulate any level or levels intermediate between lexical representation and pronunciation, or whether it is sufficient to have just these two levels with suitable (preferably universal) constraints on what a well-formed mapping between them can be. To the extent that one needs to postulate intermediate or alternative levels which are not determinate on the basis of the adult language, one is ascribing an independent system to the child. To the extent that one is able to predict the child's pronunciation on the basis of the adult language and general learning strategies, one is denying that the child has his own system. At a tangent to this issue I want to say a little about the putatively innate status of such principles in the event that there are any.

Two-year-old children typically pronounce words of the adult language as in (1):[2]

(1) *duck* → gʌk *blue* → bu:
pin → bin *bottle* → bɔkəl

The reason for such 'mispronunciations' has been variously given as the child's immature perceptual ability (Braine, Locke, Macken, Timm, Waterson, etc.), his motor inability or disinclination to produce (some of) the sounds involved (Ferguson and Farwell, Pupier, etc.), the effect of innate neutralization rules that the child has to 'unlearn' (Cairns, Stampe), the effect of realization rules conforming to certain universal constraints (Smith), the effect of idiosyncratic output constraints (Ingram, Menn, Priestley), or some combination of these (nearly everybody).

Although probably no specific individual would subscribe to either of them, one can represent the range of alternatives by reference to a maximal and a minimal model of child phonology, given in (2a) and (2b) respectively:[3]



In these diagrams boxes represent definable levels of representation linked by sets of formally similar rules or principles indicated by labelled arrows. The boxes might also be labelled with the names of their hopeful sponsors: thus, in (2a) the level of 'Organizational representation' might be called 'David Ingram', the 'Output lexicon' might be called 'Lise Menn' and so on. However, I shall shroud them in decent anonymity for the present.

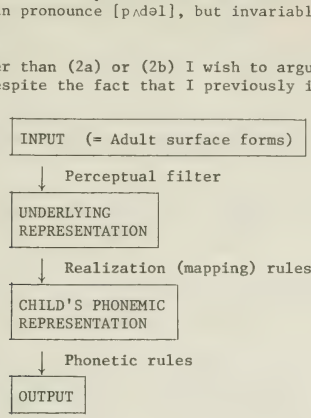
To see what is being claimed let us take a single simple example: the pronunciation [bɔkəl] for '*bottle*'. If perceptual factors are responsible for this deviation then the child's lexical representation for the word would be something close to his output - [bɔkəl]. If productional (motor) difficulties were responsible then the lexical representation would be closer to or identical with the adult /bɔtəl/. If organizational factors were responsible then it would be being claimed that the lexical representation was 'correct' (i.e. equivalent to the adult surface form) and that there was furthermore no motoric difficulty inherent in the pronunciation of the sequence [b+ɔ+t+ə+l], but that for independent psychological reasons the child chooses to set up an additional level of representation which disallows sequences of the relevant form. That this is not a priori ridiculous (though I suspect a posteriori that it is wrong) can be seen from such examples as (3):

(3) *puddle* → pʌgəl
puzzle → pʌdəl

where the child can perceive the difference between *puddle* and *puzzle* correctly, can pronounce [pʌdəl], but invariably reproduced '*puddle*' as [pʌgəl].

Rather than (2a) or (2b) I wish to argue for the position represented in (4), despite the fact that I previously invested in (2b).

(4)



(4) embodies the claim that perceptual factors are crucially involved in the setting up of (a small subset of) underlying representations for the child, but that there is a single level of lexical representation and a single set of phonological realization rules (also known as 'incompetence' rules (Smith 1971)) which map this representation onto a level describable as the child's phonemic system: i.e. the level immediately prior to the operation of low-level phonetic (allophonic) rules.[4]

Let us look first at the child's perceptual ability and its implications for his lexical representation. It is uncontroversial that the child's perception is somewhat in advance of his productive ability; what is not clear is whether any aspect of his infantile pronunciation needs to be attributed to perceptual difficulty. On the basis of various kinds of evidence I postulated a few years ago that the brief answer to this question was 'no'. Three representative pieces of such evidence are summarized here. First, the child could successfully discriminate in decontextualized situations pairs of words such as those in (5) which he produced as homophones.

(5) *bad*, *bat* → bæt
teddy, *Daddy* → dædi:
mouth, *mouse* → maus

Second, he typically implemented changes to his system 'across-the-board' as indicated in (6). That is, having learned to produce a particular configuration for one word: e.g. [blu:] for *blue* in place of the earlier [bu:], this new-found ability was immediately generalized to all and only those words containing a postconsonantal liquid in the adult language,

indicating that these words must have been correctly stored.

(6)

t_1			t_2			$*t_i$		
blue	→	bu:	*blue*	→	blu:	*blue*	→	blu:
play	→	bei	*play*	→	plei	*play*	→	bei

Crucially there was no stage t_i where some of those words changed and some did not. Third, he formed plurals differentially for words which ended in different segments in the adult language but the same segment in his speech. Thus the examples in (7):

(7)

cloth	→	klɔt	*cloths*	→	klɔtid
cat	→	kæt	*cats*	→	kæt
horse	→	ɔ:t	*horses*	→	ɔ:tid

indicate that his lexical representation for these words could not be the same as his output, and that he perceived the differences in the adult words on the basis of which these contrasts are predictable.

Furthermore, the hypothesis that children's perceptual abilities are perfect seems to be consonant with experimental work on the perceptual ability of infants. Thus Eimas and his colleagues (Eimas 1975; Cutting & Eimas 1975) have demonstrated the ability of infants at one month to discriminate categorically between segments differing in voice onset time such as [pa - ba], and between [r] and [l]; Jusczyk and Thompson (1978) demonstrated the ability of two-month-olds to discriminate place of articulation among [b - d - g] and stress contrasts of the sort [dága - dagá]; Morse (1972) that by six weeks they can discriminate falling and rising pitch; etc. Even the subsequent discoveries that the infant's categorical perception of speech was shared by, of all things, the chinchilla (Kuhl & Miller, 1975); that nonspeech sounds were also categorically perceived by both infants (Jusczyk et al. 1977) and adult musicians (Siegel & Siegel 1977), seemed only to indicate that the human, or perhaps mammalian, perceptual ability was up to anything language or the world in general could throw at it. (Eilers 1978 provides a good recent survey.) Moreover work by Barton (1976) showed that this infantile ability persists beyond the first few months as children can still make such distinctions (specifically the voicing contrast) at the advanced age of 1½ to 2 years. However, Barton also noted that the two-year-old's perceptual ability is dependent on the familiarity of the words heard; so *coat* and *goat* might be discriminated whereas *coast* and *ghost* might not be. This casts doubt on the generality or consistency of the child's perception and recent elegant work by Macken (1978a) has convinced me that I was wrong: there is a specifiable class of cases where the child's mispronunciations must be attributed to perceptual mistakes.

Following up an observation of Braine's (1976) that some of the child's mispronunciations 'cried out for a perceptual explanation', Macken has conclusively demonstrated that 'the child does have (in some cases) a unique, non-adult underlying representation . . . one that is, moreover, based on a perceptual error'. The argument is based on a difference between two kinds of free variation: intra-word free variation of the sort exemplified by the alternative pronunciations of a single word, as in (8):

(8) *frog* → flɔg, wɔg, wlɔg, βrɔg

and interword free variation of the sort exemplified in (9);

(9) *bottle* → bɔkəl *cuddle* → kʌdəl
puddle → pʌgəl *beetle* → bi:təl
kennel → kɛŋəl

where a widespread rule which applies regularly in some words to velarise a coronal stop before a dark [ɫ], has apparently arbitrary exceptions. Whereas the former kind of free variation can be satisfactorily accounted for on the assumption that the child has the adult form as his lexical representation and mangles it by rules some of which are optional, no such explanation can account for the latter type (interword free variation) where the pronunciation of each word is selfconsistent, which presumably resides precisely in a difference in the child's lexical representation. But if the cause of the variation is a difference in the lexical representation of the two items (e.g. *puddle* and *cuddle*) then it cannot be the case that both of them are represented in terms of the adult surface form. This establishes that the child does not always internalize the adult form; that the cause of the difference is perceptual follows from two further observations: the piece-meal nature of the changes over time in such exceptional items, and the fact that when the perceptual encoding rule is changed it is applied incorrectly to some items. Whereas the vast majority of mapping rules were virtually exceptionless, some 20% of those items which should have undergone the velarisation rule converting *puddle* to [pʌgəl] failed to do so. More importantly, some items, exemplified in (10), acquired their correct pronunciation in a gradual, piece-meal fashion rather than simultaneously across-the-board; indicating that the child was getting things right word by word and not by phonological generalization.

(10)

	t_1		t_2		t_3
bottle	→ bɔkəl	*bottle*	→ bɔkəl	*bottle*	→ bɔtəl
puddle	→ pʌgəl	*puddle*	→ pʌgəl	*puddle*	→ pʌdəl
cuddle	→ kʌdəl	*cuddle*	→ kʌdəl	*cuddle*	→ kʌdəl
beetle	→ bi:kəl	*beetle*	→ bi:təl	*beetle*	→ bi:təl

More significant still, there were some adult words, listed in (11),

(11)

	t_1		t_2		t_3
circle	→ tə:kəl	*circle*	→ sə:kəl	*circle*	→ sə:təl
pickle	→ pikəl	*pickle*	→ pikəl	*pickle*	→ pitəl
winkle	→ winkəl	*winkle*	→ winkəl	*winkle*	→ wintəl

with an original velar, which were incorrectly given an alveolar at the time the examples in (10) were changing. That is, after being correctly pronounced as, for instance, [tə:kəl] for many months, *circle* and similar words acquired an anomalous, incorrect, pronunciation with [t]. This is a clear indication that the child's lexical representation was vacillating: as a result of not discriminating perceptually between /t/ and /k/ in this environment, he had to learn piece-meal which sound went with which word after he had made the relevant perceptual progress. I had earlier stated

categorically that such mistakes were impossible and would constitute counter-exemplary data to my model of phonological acquisition (roughly equivalent to (2b), cf. Smith 1973:ch. 4). Macken has proved that they are not only possible but occurred in the data on the basis of which I had made my original prediction. My only solace is that I have demonstrated intellectual honesty in the sense of Lakatos (1970), who wrote that: 'intellectual honesty consists in stating the precise conditions under which one will give up one's position.'

It should perhaps be emphasized that the proportion of words subject to perceptually induced malformation is small (exactly how small I am not sure) and that there is no evidence to my knowledge that the mispronunciations of normal children are a reflection of their limited perceptual ability in the way that the mutilated productions of dysphasic children do seem accurately to reflect a perceptual deficit. Specifically, Tallal and her colleagues have demonstrated (cf. Tallal & Stark 1978) that a group of children with delayed onset of speech were incapable of perceiving sound contrasts involving rapid temporal processing: e.g. consonants with their characteristic formant transitions as opposed to steady state vowels.

Nevertheless the minimal system represented by (2b) is clearly wrong. It remains to be seen whether there is any justification for any other part of the baroque complexity of (2a). To decide this matter we need to disentangle a number of different strands; most importantly the difference between a model containing say (2 + n) different levels of representation and a model containing just two levels of representation but n different kinds of motivation for the rules linking those levels. If we can decide this matter then we shall also be making progress on the question of whether children 'have their own system' or not. That is, in (4) there is a single mapping from the underlying representation to the surface: i.e. from something approximating very closely to the adult language to the child's pronunciation (but cf. footnotes 3 and 4). In (2a) on the other hand the relation between adult and child forms is much less direct and the intervening levels of representation are what constitute the child's own system (in at least one sense of this expression). It is of course trivially true that since each child is different to every other child he must have his own system in some sense; the claim that 'he has no system of his own' then must be taken to imply that all that is required to predict the kind of pronunciation the child will come up with is the underlying representation and a set of principles governing how these representations can be mapped onto the surface: i.e., the child's phonology is parasitic on the adult language and can deviate from it only according to a very limited set of strategies. To be specific, I proposed (1973) that there is a single set of rules mapping lexical representations onto the phonemic forms and that all such rules were formally and functionally constrained by universal principles. In particular, any rule had to implement one or more of the strategies serving to maximize the phenomena illustrated in (12):

(12) a) Consonant or vowel harmony as in the pronunciation [gʌk] for '*duck*'.

b) Cluster reduction as in the pronunciation [bu:] for '*blue*'.

c) Systemic simplification as in the pronunciation [mæt] neutralising the four way contrast in *'mat'*, *'mass'*, *'mash'*, *'match'*.

d) Grammatical simplification as in the pronunciation [ai] for both of the singular *'eye'* and plural *'eyes'*.

Differences between various children are then ascribable to the differential implementation of these four strategies: thus one child might implement consonant harmony by changing *duck* to [gʌk], another by changing it to [dʌt], and so on. It should perhaps be noted at this point that the rules and constraints I posited then now have a little less work to do as a result of the conclusions discussed above on the role of perception. Apart from this class of exceptions I think the constraints still stand. Their big advantage is that they exclude certain kinds of process from effecting relations between adult and child forms. For instance, regular dissimilation, as opposed to assimilation, is claimed to be impossible. Given that such dissimilation occurs both historically and in fast speech, this is a nontrivial and easily falsifiable prediction.[5]

Unfortunately the predictions made by (12) seem to be not only falsifiable but false. The clearest counterexamples are provided by so-called 'idiosyncratic strategies' of the sort described by Ferguson and Farwell (1975) and illustrated from Priestley (1977) in (13):

(13) Sundry polysyllabic words adopt the canonical form CVjaC
e.g. *basket* → bajan

Such examples, where the set of forms concerned do not constitute a phonologically well-defined class, clearly fall outside the range of my strategies and hence appear to disconfirm my position. However, Macken's re-evaluation of the role of perception leads me to hope that the constraints on realization rules can be maintained as tightly as before, with all apparent counterexamples--such as those of Priestley's--being perceptually conditioned and hence not due to the action of the realization rules at all. Although (dis)confirmation of this hypothesis can come only from perceptual experiments on children manifesting such idiosyncratic strategies, it is significant that all the convincing examples are of polysyllabic items: i.e. items which are notoriously difficult for young children to master perfectly for simple reasons of complexity, and items where there is sufficient redundancy for it to be unnecessary for the child to internalize the entirety of the word. Moreover in their longitudinal development these forms also showed the piece-meal modification expected of perceptually rather than strategically induced deformations.

I have always considered strategies of the sort listed above as a universal set of necessary and sufficient conditions on mapping rules which the child, as part of his innate endowment, brings to bear on the task of language learning. An alternative interpretation of the complex relationship between adult and child forms is that, intermediate between the underlying representation and the phonemic representation which serves as the basis for allophonic rules, are one or more further levels of representation idiosyncratic to the child. Typical examples in the literature are provided by Ingram (1974, 1976), Waterson (1971) and by the Morpheme

Structure conditions in Smith (1973) which I in fact argued against. Ingram (1974, 1976) suggests three levels of representation--perceptual, organizational and productional--such that the pronunciation of a word like *duck* as [gʌk] may be due to perceptual problems: the child cannot hear the difference between [dʌk] and [gʌk]; to productional difficulties: he cannot pronounce [dʌk]; or to organizational constraints: he produces all words with consonant harmony for independent reasons. Unfortunately Ingram presents no evidence in favor of his organizational level except its logical possibility. There is clearly a difference between providing a teleology for the constraints that exist and setting up a linguistic level for each kind of explanation. I am not denying that some of the divergencies between adult and child forms are ultimately explicable in terms of the organizing principles the child brings to bear on language learning--far from it--what I am denying is that there is evidence for setting up a further level of representation. Given universal constraints on mapping rules and Occam's razor one has no choice but to do without such a level.

On the same lines as Ingram, Menn (1978a) suggests that many though not all 'rules are best seen in terms of the satisfaction of output constraints' of the sort 'no consonant clusters are allowed'. That is, Menn is claiming that the absence of consonant clusters is to be attributed not (or not only) to the operation of a strategy defined over the adult input, but wholly or in part to the child's conforming to a template peculiar to his own system. It is then a clear instance of a claim that the child does have his own system.

Before attempting to evaluate this claim two points need to be made: first, that realization rules (i.e. mapping rules parasitic on the adult language) are necessary anyway so that the correct output is associated with the correct lexical representation; second, that in many, perhaps most, instances realization rules and output constraints make equivalent predictions. Accordingly Occam makes it incumbent on an output constrainer to justify invoking the extra construct.

Except for the case of 'idiosyncratic strategies' which I have already suggested are perceptual in nature, I know of no strong evidence (but cf. note 6) in favor of a level such as the 'Output lexicon' in (2a), and I believe that there are objections to it in principle. Output constraints are redundant. To justify this claim let me first give a case where a mapping rule can handle the data but an output constraint is unable even to describe what is going on. For somewhat over a year A had a rule converting adult sequences of the form $C_{\omega}VC$ to the form CV_{p}, resulting inter alia in the examples in (14):

(14) *quick* → kip
quite → kaip
twice → taip

Given that there was another rule reducing clusters of a consonant followed by a sonorant to the consonant, as exemplified in (15):

(15) *green* → gi:n

black → bæk
new → nu:

the expected output of forms with postconsonantal /w/ would be CVC. Given further that CVC forms with all consonants in position C_2 were allowed, as witness the examples in (16):

(16) *skip* → kip
kiss → kit
kick → kik

then the forms given in (14) can clearly not be accounted for by an output constraint. However, a strategy of consonant harmony, assimilating the final consonant to the point of articulation of the labial /w/, accounts for the facts simply. Realization rules are then clearly necessary merely to achieve observational adequacy. Moreover, the relation between the segments in the underlying and the segments in the child's surface forms is a many:many one--a situation which output constraints are unable to handle. Although the neutralization of an adult contrast by the child (of the sort seen in (12c)) can be described by either output constraints or realization rules, the reverse situation, where a single segment in the adult language is realized alternatively as two or more different segments by the child, is captured very simply by means of an optional rule, but cannot be handled by output constraints. The notion 'optional output constraint' is incoherent.

Further, consider the longitudinal development of certain forms in A's speech, as illustrated in (17):

(17)

	t_1		t_2		t_3	
slip	→	lip	*slip* →	lip	*slip* →	ɬip
bit	→	bit	*bit* →	bit	*bit* →	bit
pit	→	bit	*pit* →	pit	*pit* →	pit
spit	→	bit	*spit* →	pit	*spit* →	pit

At time t_1 there were two relevant processes operative: one deleting preconsonantal /s/, the other neutralizing the adult voicing contrast. At t_2 the rule neutralizing voicing disappeared and the child correctly contrasted *pit* and *bit*. As *spit* still fell together with *pit* rather than *bit* at this stage, I assume that the child had a rule devoicing nonsonorant consonants after /s/ which applied before /s/ deletion; cf. (18).

(18) $[\begin{smallmatrix} C \\ -son \end{smallmatrix}]$ [-voiced] / s ___

At t_3 the child suddenly acquired a series of so-called voiceless sonorants, as illustrated in (19):

(19) *Smith* → m̥it
snail → n̥eil
slip → ɬip

That is, he regularly and consistently used sounds which, although common in Welsh, Burmese, and Classical Aztec, are not typical of British English.

It is hard to see what account of this could be given by output constraints, but these data are an automatic consequence of a process of realization rule simplification: the structural description of the rule (18) has been simplified by the omission of the feature sonorant to give the rule in (20):

(20) C → [-voiced] / s ___

It is of course impossible to prove that something does not exist: evidence for it may turn up tomorrow. I think though that output constraints come into the category of objects for whose existence no evidence has yet been given.[6]

Indirect support for this position also comes from perceptual evidence. I take it that the levels of 'Underlying representation' and 'Child's phonemic representation' in (4) have psychological reality for the child (as also do the rules linking them, as will be clear from the preceding discussion of rule simplification), though they have reality of somewhat different kinds. Thus the child's perception is in terms of the system given by the former level and his production in terms of the system defined by the latter level. Dodd (1975) has shown in her experimental work that children are better at understanding an unfamiliar adult's tape-recorded speech than they are at understanding tape-recordings of their own speech. Their understanding of their own speech was successful precisely to the degree that it approximated to the adult forms: i.e. the greater the divergence the less the understanding. Their understanding adult forms which they often did not differentiate in their own speech clearly implies that they have the adult form as one part of their lexical representation. Their failure to understand their own forms equally clearly detracts from the plausibility of the argument that that form too is stored in any way in their heads. This supports the claim for the psychological reality of the level of lexical representation in (4); the reality of the level of the child's phonemic representation is attested to by simple productional facts (cf. Braine 1976). It is not clear to me what, if anything, could constitute evidence for the psychological reality of a further level.

It might also be added that postulating this extra level, which is an addition to the elegant Kiparsky and Menn (1978) model of language acquisition, lands Menn (1978a) in further descriptive problems which necessitate building the Duke of York gambit (cf. Pullum 1976) into auto-segmental phonology.[7] I will not pursue this horror story here. Again I should emphasize that I do not doubt the heterogeneity of the explanations for the rules in (4): the examples in (9) may well be explained ultimately in perceptual terms, the examples in (14) in organizational terms, those in (17) in terms of the psychological reality of the mapping rules, and probably some changes--e.g. those in (12c)--are due solely to the motor difficulty of the articulations involved. None of these, however, motivates the setting up of further psychologically real linguistic levels.

Another way in which the child's lexical representation might diverge from the adult surface form is in the completeness of the specification of segments and sequences of segments. It has been suggested by, among

others, Wilbur (forthcoming) and Ingram (1974) that in those cases where the child's pronunciation diverges from the adult's by the absence in the former of a segment in the latter, the child has a partial specification of the word as his lexical representation rather than the complete adult surface form. Despite the fact that I was party to a somewhat similar suggestion myself once (cf. Morton and Smith 1974) the claim seems implausible. The claim is that if a child pronounces *spoon* as [bu] for instance, then his representation of it may be taken to be X*pu*X, where X indicates the presence of a segment or segments either without further specification or with incomplete specification (Ingram 1974). However, partial specification (with the limiting case being that of noise indeterminate from all else except silence) is incompatible with what we know about perception in two-year-olds. Thus any minimal pair of the adult language whose minimal difference resides in a segment or segments represented by Ingram as X is predicted as equivalent to the child. Either observed across-the-board change or perceptual testing of minimal pairs of the sort conducted by Barton would choose, for any child, between the total and the partial hypothesis. Further evidence against the partial specification hypothesis comes from the reported avoidance by some children of homonymous outputs. That is, otherwise regular rules have exceptions if and only if the failure of some form to undergo such a rule avoids the child's producing homophones.

The only area where specifications might be incomplete is again where there is perceptual confusion 'upstream' of the lexical representation. If such forms are represented partially then either of two or more neutralized adult equivalents should surface in the child's production. If such forms are represented totally but incorrectly then only one form should appear until the mistake is discovered and rectified.

Earlier I suggested that the constraints on realization rules I have been defending are universal. If true, the most plausible, albeit not necessary, assumption is that this universality reflects innateness: it is biologically determined. Innateness is a dangerous area, so let me hedge (a little). Kiparsky and Menn (1978) have argued convincingly that the rules typical of child phonology are all invented by the child as solutions to specific problems, rather than being innately specified. While I accept this in one sense (it is fully compatible with Smith 1973), and take it as a cogent critique of Stampe's (1969) position, it still seems necessary to account in more general terms for the rather limited range of solutions that children do come up with. It is at this level of abstraction that it makes sense to talk about innateness. Note, incidentally, that the term 'innate' itself has been given a number of different interpretations. Following Aslin and Pisoni (1978--in turn following Gottlieb) we can distinguish among:

> i) an ability which is fully developed at birth but which may be lost if external stimulation is absent. A typical example is provided by vision;
> ii) an ability which is partially developed at birth such as birdsong (in birds that is);
> iii) an ability which is undeveloped at birth but which is triggered off by some external stimulus, either immediately as in

imprinting or after further maturation.

It seems clear that the ability to invent phonological rules is an instance of this final type. The ability is not present at birth; indeed, it may not operate fully in the earliest stages of language acquisition, as witness the existence of 'phonological idioms' which by-pass the rules: e.g. the early pronunciation [priti] for *pretty* reported by Leopold (1939-47), and it is obviously dependent on external stimulation. Note that this is not a claim for empiricist inductivism: rather it is a claim that the hypotheses the child can come up with to solve his problems belong to an antecedently well-defined set, constrained by the strategies discussed above and illustrated in (12).

Some corroboration for this view comes from the phonological development of the prelinguistically profoundly deaf (cf. Dodd & Hermelin 1977) where the external stimulus is not even auditory. Despite their disadvantages, deaf children internalize phonological rules which are in major respects the same as those of the hearing: i.e. their deformations from the adult norm conform to the same constraints as do those of the hearing (and as those of some dysphasics do not--cf. Smith 1975). Indeed, it also turns out that much of the information encoded in speech can be inferred by hearing listeners on the basis of visual (lip-read) input (Dodd 1977). That is, the hypotheses the child can come up with are not only 'pre-ordained' but are also modality neutral.

Elsewhere in this volume, McCawley and Comrie are both severely critical of the postulation of universals either on the basis of very limited (English language) data or inherent implausibility. The answer to their strictures resides in proposing specific, motivated, falsifiable universal claims. Although my examples are drawn only from English and largely from one child, I have tried to do just that.

NOTES

*I would like to dedicate this article to John Trim, who first taught me linguistics and who first encouraged me to work on the acquisition of phonology.

[1]I am grateful to Deirdre Wilson for her comments on a preliminary version of this paper, to Charles Ferguson and John Locke for a number of stimulating discussions after the lecture and to Barbara Dodd and Geoff Pullum for helpful observations on the revised version. The foregoing are to be taken as agreeing only with those parts of the article which are true.

[2]Unless otherwise indicated acquisitional data are taken from the analysis of Amahl in Smith (1973). For exegetical clarity I have omitted irrelevant phonetic detail, and in some cases have included examples from different stages of A's development without comment.

[3]In fact both these diagrams embody an important oversimplification

in that they make no distinction between the phonetic and the phonemic shape of the adult surface forms. I have assumed tacitly that the input is in some kind of phonemic code, but the longitudinal development of the acquisition of some contrasts could be taken to argue against such an assumption. For instance, A acquired the voiced/voiceless contrast first in final position, then medially and last of all in initial position (cf. Smith 1973:95-96). It may either be the case that this is the result of the child not identifying the initial, medial and final allophones of the adult phonemes; or it may be that, despite such identification, the different positions are more salient for him and hence override the phonemic identity. It is not clear to me what kind of data would decide between these alternatives. In those cases where all and only the allophones of one adult phoneme change their realization simultaneously we do have evidence for the identification by the child of the phonemic status of the allophones. Unfortunately the necessary stipulation 'all and only' makes such cases extremely rare. Note that this issue is separate, or at least separable, from the issue as to whether the child stores lexical items in a phonemic or a phonetic code (cf. Morton and Smith 1974 for a brief discussion, and also note 4 below).

[4]Mutatis mutandis the remarks of note 3 apply here too. That is, it is hard to find evidence which demonstrates conclusively that the level designated 'Child's phonemic representation' is psychologically real and distinct from the phonetic representation of the 'Output'.

[5]It follows that Ingram's attempt (1974) to reanalyze some of the data from Smith (1973) in terms of dissimilation is a priori undesirable. In fact Ingram provides no justification for his analysis, and to replace a constrained system by a relatively unconstrained system without arguing for the increase in expressive power entailed thereby is counterproductive.

[6]Evidence for output constraints could be found in principle if it could be shown that two rules, justified by reference to different strategies, conspired to produce the same output. Macken (1978b) has suggested that one such example obtains in the interaction of consonant harmony and metathesis, which conspire to produce the constraint that coronals may not appear before labials or velars (cf. Morpheme Structure Condition 4 in Smith 1973:41-42). Such a condition, captured by a Morpheme structure rule or an output constraint, appears to encapsulate a generalization which is missed by the two realization rules. However, whether this example finally turns out to substantiate output constraints or not will in turn depend on an experimental evaluation of Berman's (1977) claim that all metatheses are perceptually conditioned. I suspect that they are not, but if not, then the case for output constraints is correspondingly strengthened, because of the implication that consonant harmony and metathesis are processes of the same, realizational, type.

[7]To be specific, Menn's position entails that a word such as *fish* pronounced as, say, [iʃ] by the child has the following derivation. It starts with a full specification of the content and linear sequence of the three segments /f/, /i/, and /ʃ/. Then, because fricatives occur in the child's speech only in final position, the redundant linear sequence of the segments is suppressed in his lexical representation. Finally,

to account for the child's pronunciation with [ʃ] in final position, linear ordering is reinstated. Ptolemy would have approved.

REFERENCES

Aslin, R. and D. Pisoni. 1978. Some developmental processes in speech perception. Paper presented at the conference on Child Phonology, Bethesda, Md. May, 1978.

Barton, D. 1976. The role of perception in the acquisition of phonology. University of London Ph.D. thesis. Relevant parts of this appear in Phonemic discrimination and the knowledge of words in children under 3 years. Papers and reports on child language development, Stanford.

Berman, R. 1977. Natural phonological processes at the one-word stage. Lingua 43. 1-21.

Braine, M. 1974. On what might constitute learnable phonology. Lg. 50. 270-99.

------. 1976. Review of Smith (1973). Lg. 52. 489-98.

Cairns, C. 1969. Markedness, neutralisation and universal redundancy rules. Lg. 45. 863-815.

Cutting, J. and P. Eimas. 1975. Phonetic feature analysers and the processing of speech in infants. The role of speech in language, ed. by J. Kavanagh and J. Cutting, 127-48. Cambridge, Mass.: M.I.T. Press.

Dodd, B. 1975. Children's understanding of their own phonological forms. Quarterly Journal of Experimental Psychology 27. 165-72.

------. 1977. The role of vision in the perception of speech. Perception 6. 31-40.

------ and B. Hermelin. 1977. Phonological coding by the prelinguistically deaf. Perception and Psychophysics 21. 413-17.

Eilers, R. 1978. The complex nature of infant speech perception. Paper presented at the conference on Child Phonology, Bethesda, Md. May, 1978.

Eimas, P. 1975. Auditory and phonetic coding of the cues for speech: Discrimination of the r - l distinction by young infants. Perception and Psychophysics 18. 341-47.

Ferguson, C. and C. Farwell. 1975. Words and sounds in early language acquisition. Lg. 51. 419-39.

Ingram, D. 1974. Phonological rules in young children. Journal of Child Language 1. 49-64.

------. 1976. Phonological analysis of a child. Glossa 10. 3-27.

Jusczyk, P., B. Rosner, J. Cutting, C. Foard and L. Smith. 1977. Categorical perception of non-speech sounds by two-month-old infants. Perception and Psychophysics 21. 50-54.

Jusczyk, P. and E. Thompson. 1978. Perception of a phonetic contrast in multisyllabic utterances by two-month-old infants. Perception and Psychophysics 23. 105-9.

Kiparsky, P. and L. Menn. 1978. On the acquisition of phonology. Language and thought, ed. by J. McNamara, 43-78. New York: Academic Press.

Kuhl, P. and J. Miller. 1975. Speech perception by the chinchilla: Voiced - voiceless distinction in alveolar plosive consonants. Science 190. 69-72.

Lakatos, I. 1970. Falsification and the methodology of scientific research programmes. Criticism and the growth of knowledge, ed. by I. Lakatos and A. Musgrave, 91-196. Cambridge: Cambridge University Press.
Leopold, W. 1939-47. Speech development of a bilingual child: A linguist's record. 4 vols. Evanston, Ill.: Northwestern University Press.
Locke, J. 1979. The child's processing of phonology. Minnesota symposium on child psychology 12, ed. by W. Collins, 83-119. Princeton, N.J.: Lawrence Erlbaum.
Macken, M. 1978a. The child's lexical representation: Evidence from the "*puzzle-puddle-pickle*" phenomenon. Mimeo. Stanford University.
------. 1978b. Aspects of the acquisition of stop systems: A cross-linguistic perspective. Paper presented at the conference on Child Phonology. Bethesda, Md. May 1978.
Menn, L. 1978a. Phonological units in beginning speech. Proceedings of the symposium on segement organization and the syllable, ed. by A. Bell and J. Hooper.
------. 1978b. Phonological theory and child phonology. Paper presented at the conference on Child Phonology. Bethesda, Md. May, 1978.
Morse, P. 1972. The discrimination of speech and non-speech stimuli in early childhood. Journal of Experimental Child Psychology. 14. 477-92.
Morton, J. and N. Smith. 1974. Some ideas concerning the acquisition of phonology. Problèmes actuels en psycholinguistique. Editions du C.N.R.S. 161-76.
Priestley, T. 1977. One idiosyncratic strategy in the acquisition of phonology. Journal of Child Language 4. 45-65.
Pullum, G. 1976. The Duke of York gambit. Journal of Linguistics 12. 83-102.
Pupier, P. 1977. Quelques observations sur l'acquisition de la phonologie par des enfants montréalais de 2 ans. Montreal Working Papers in Linguistics 9. 175-202.
Siegel, J. and W. Siegel. 1977. Absolute identification of notes and intervals by musicians. Perception and Psychophysics 21. 143-52.
Smith, N. 1971. Puggles and lellow lollies. Listener 86. No. 2227. 759-60.
------. 1973. The acquisition of phonology: A case study. Cambridge: Cambridge University Press.
------. 1975. Universal tendencies in the child's acquisition of phonology. Language, cognitive deficits and retardation, ed. by N. O'Connor. IRMMH Study Group 7. Woburn, Mass.: Butterworths.
Stampe, D. 1969. The acquisition of phonetic representations. Chicago Linguistic Society 5. 443-54.
Tallal, P. and R. Stark. 1978. Do the speech production abilities of language impaired children mirror their perceptual abilities? Paper presented at the Institute of Acoustics symposium on Children's Speech: Production and Perception, U.C.L. June, 1978.
Timm, L. 1977. A child's acquisition of Russian phonology. Journal of Child Language 4. 329-39.
Waterson, N. 1971. Child phonology: A prosodic view. Journal of Linguistics 7. 179-211.
Wilbur, R. (forthcoming) Theoretical phonology and child phonology. Phonology in the seventies, ed. by D. Goyvaerts.

The following special issues are in preparation:

Relational Grammar and Semantics
Editor: Jerry L. Morgan

Studies in Arabic Linguistics
Editor: Michael J. Kenstowicz

Studies in Language Variation: Nonwestern Case Studies
Editor: Braj B. Kachru

South Asian Linguistics: Syntax and Semantics
Editor: Yamuna Kachru

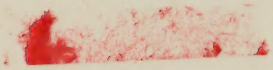
STUDIES IN THE LINGUISTIC SCIENCES

The following issues of SLS are out of print:
Spring 1971, Vol. 1, No. 1, *Papers in General Linguistics;* Fall 1971, Vol. 1, No. 2, *Papers on Hindi Syntax,* Editor: Yamuna Kachru; Spring 1972, Vol. 2, No. 1, *Papers on Syntax and Semantics,* Editor: Georgia M. Green; Fall 1972, Vol. 2, No. 2, *Studies in Baltic Linguistics,* Editors: Hans H. Hock and Michael J. Kenstowicz; Spring 1973, Vol. 3, No. 1, *Papers in General Linguistics;* Fall 1973, Vol. 3, No. 2, *Papers on South Asian Linguistics,* Editor: Braj B. Kachru; and Spring 1976, Vol. 6, No. 1, *Papers in General Linguistics,* Special section: *Topics in relational grammar,* Editor: Jerry L. Morgan.

The following issues are available:

Spring 1974	Vol. 4, No. 1, *Papers in General Linguistics*	$3.50
Fall 1974	Vol. 4, No. 2, *Papers on Phonetics and Phonology* Editors: Charles W. Kisseberth and Chin-W. Kim	$3.50
Spring 1975	Vol. 5, No. 1, *Papers in General Linguistics*	$3.50
Fall 1975	Vol. 2, No. 2, *Papers on Historical Linguistics: Theory and Metho* Editors: Ladislav Zgusta and Hans H. Hock	$3.50
Fall 1976	Vol. 6, No. 2, *Papers on African Linguistics* Editors: Eyamba G. Bokamba and Charles W. Kisseberth	$3.50
Spring 1977	Vol. 7, No. 1, *Papers in General Linguistics*	$3.50
Fall 1977	Vol. 7, No. 2, *Studies in East Asian Linguistics* Editors: Chin-chuan Cheng and Chin-W. Kim	$3.50
Spring 1978	Vol. 8, No. 1, *Papers in General Linguistics*	$5.00
Fall 1978	Vol. 8, No. 2, *Linguistics in the Seventies: Directions and Prospects* Editor: Braj B. Kachru	$10.00

Orders should be sent to:
SLS Subscriptions, Department of Linguistics
4088 Foreign Languages Building
University of Illinois
Urbana, Illinois 61801




BOUND BY
HNM

UNIVERSITY OF ILLINOIS-URBANA
410ST92 C001
STUDIES IN THE LINGUISTIC SCIENCES; URBA
8 1978
3 0112 025054617